



Standing up or standing by: Bystander intervention in cyberbullying on social media

new media & society

2021, Vol. 23(6) 1379–1397

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1461444820902541

journals.sagepub.com/home/nms**Sai Wang** 

City University of Hong Kong, Hong Kong

Abstract

Cyberbullying often happens in the presence of bystanders whose behaviors play a key role in changing dynamics of the situation. To examine the factors influencing cyberbystander likelihood of intervening in cyberbullying on social media, a 2 (degree of cyberbullying severity: high vs low) $\times$ 2 (level of interpersonal similarity: high vs low) between-subjects experiment was conducted ($N=132$). Results indicated that cyberbystanders' empathic responses depended on the severity of cyberbullying. The more empathy bystanders felt, the more likely they were to publicly and privately intervene. Moreover, the indirect effect of cyberbullying severity on willingness to intervene through empathy was more pronounced when the interpersonal similarity between victim and cyberbystander was low, rather than it was high.

Keywords

Cyberbullying, cyberbystander, empathy, intervention, severity, social media

While the rise of information and communication technologies (ICTs) has provided great opportunities for people to access information and stay connected with others, it also increases the risk of online users being confronted with threatening situations, such as cyberbullying. As an extension of offline bullying behaviors undertaken in a new modality, cyberbullying refers to “the behavior performed through electronic or digital media by individuals or groups that repeatedly communicates hostile or aggressive messages intended to inflict harm or discomfort on others” (Tokunaga, 2010: 278). In general,

Corresponding author:

Sai Wang, Department of Media and Communication, City University of Hong Kong, M5013, Run Run Shaw Creative Media Center, 18 Tat Hong Avenue, Kowloon, Hong Kong.

Email: saiwang3-c@my.cityu.edu.hk

cyberbullying serves as an umbrella term that encompasses similar constructs such as electronic bullying (Kowalski and Limber, 2007) or online harassment (Ybarra and Mitchell, 2004). Cyberbullying has detrimental impacts on victim's physical and psychological well-being. Victims tend to feel depressed and powerless (Anderson et al., 2014), and some of them even attempt suicide (Patchin and Hinduja, 2012).

Similar to offline bullying scenario, individuals involved in a cyberbullying situation can be classified into three major roles: perpetrators who initiate the bullying, victims who are afflicted with insult, and bystanders who witness the incident (called "cyberbystanders"). In the context of cyberspace, perpetrators often willfully aggress against a victim under the cover of online anonymity, whereas victims are less able to defend themselves. Sometimes, the impulse for retaliation results in the reversal of perpetrator-victim roles by turning the victim into the perpetrator (Kowalski and Limber, 2007). Cyberbystanders, on the other hand, represent third-party participants whose reactive behaviors are crucial to influence the course of cyberbullying and its consequences. For instance, bystander intervention (e.g. telling a perpetrator to stop, defending a victim) is effective in terminating a bullying episode and mitigating negative effects on victims (Salmivalli, 2010). Conversely, bystanders' escalating behaviors (e.g. assisting a perpetrator to assault a victim), or even passive observation, result in increased severity of incidents (O'Connell et al., 1999). Similarly, public intervention of cyberbystanders motivates other onlookers to do likewise by challenging the bully or defending the victim (Anderson et al., 2014). In doing so, they may threaten the status of perpetrators and stop their bullying behaviors, thus alleviating deleterious consequences of cyberbullying on victims. Despite potential influence, most cyberbystanders tend to remain outsiders when witnessing a cyberbullying incident, while few are willing to step in and help (Dillon and Bushman, 2015; Freis and Gurung, 2013; Huang and Chou, 2010). Thus, there is a great need to explore bystander intervention tactics online.

Although social media is the most common scene of cyberbullying compared to other media platforms (Chen et al., 2017; Smith et al., 2008), few studies have investigated potential factors influencing cyberbystanders' willingness to intervene in a cyberbullying situation which they come across on social media. Unlike previous research on cyberbystander behavior that has primarily built on Latané and Darley's (1970) bystander intervention model (e.g. Brody and Vangelisti, 2016; Dillon and Bushman, 2015; Kazerooni et al., 2018; Obermaier et al., 2016), this study provides a unique theoretical approach to understand how cyberbystanders respond to acts of cyberbullying. Drawing on the arousal: cost-reward model (Dovidio et al., 1991; Piliavin et al., 1981), this study investigates the effects of cyberbullying severity and interpersonal similarity on cyberbystander intervention in cyberbullying on social media. To elucidate the psychological mechanisms underlying such effects, empathic responsiveness to situation is also examined.

Theoretical background

Cyberbystander intervention in public and private dynamics

Bystander intervention can be distinguished into various forms, depending on whom an action is targeted at (i.e. directedness) or how visible an action is to others (i.e. publicness).

Latané and Darley (1970) defined direct intervention as attempting to cope with the situation straightforwardly and indirect intervention as seeking help from someone outside of the situation. Direct intervention in offline situations (e.g. breaking up a fight, rushing into a fire scene to rescue a trapped person) often involves more risks and requires more time, effort, and resources. Indirect intervention (e.g. calling the police to report a witnessed fight or fire accident), by contrast, “usually does not require a great deal of skill, strength or courage to carry it out” (Latané and Darley, 1970: 35).

In a similar vein, cyberbystander intervening behaviors can also be differentiated between direct (e.g. telling the perpetrator to stop in a comment) and indirect (e.g. reporting the incident to a website moderator) forms. Past work has indicated that cyberbystanders are more likely to take indirect forms of intervention when dealing with a cyberbullying situation than direct forms (Dillon and Bushman, 2015; Freis and Gurung, 2013). An experimental study by Kazerooni et al. (2018) found that the willingness to directly intervene was contingent on the number of bullies. When encountering a cyberbullying incident involving multiple attackers, cyberbystanders tend to engage in direct, rather than indirect, intervention. These findings provide empirical evidence for the divergent patterns of cyberbystander intervention strategies from the angle of directedness.

Besides, cyberbystander intervention can also occur either privately or publicly. Publicness refers to “the probability that one’s behavior will be observed by others or the number of others who might learn about the behavior” (Leary and Kowalski, 1990: 38). The development of communication technologies has afforded users to adjust the visibility of their messages on social media. A post that is visible to a large audience has a higher level of publicness than the same post only accessible to a small group of people. Unlike offline bullying scenarios where individuals involved in are physically co-present and bystander reactions are highly visible to others, social media provide contexts in which cyberbystanders can offer help in either a private or public manner. For example, when seeing an insulting post directed at someone on Facebook, a cyberbystander can send a private message asking the poster to stop being mean, or publish the same content through a wall post which is visible to “friends” of both the profile owner and the cyberbystander. However, the public and private patterns of intervening behavior still remain understudied. To fill this gap in existing literature, this study investigates how cyberbystanders halt cyberbullying on social media by disentangling intervention strategies from the dimension of publicness.

The arousal: cost-reward model of helping behavior in cyberbullying

The arousal: cost-reward model (Dovidio et al., 1991; Piliavin et al., 1981) has been widely used to predict bystander helping behavior in offline emergencies (e.g. Fischer et al., 2006). This model illustrates two factors influencing helping behavior. Arousal, as the first factor, describes the emotional response to the distress of others in need, which is considered as the fundamental motivational construct of the model (Dovidio, 1984). Arousal is associated with a defensive reaction through the activation of strong sympathetic nervous system (Lynn, 1966). Observation of an emergency elicits arousal characterized by aversive and unpleasant emotions, and bystanders are therefore motivated to reduce it (Piliavin et al., 1981). The second factor is cost-reward component, involving

the cognitive processes of weighing anticipated costs and rewards related with action and inaction. There are two types of potential costs for a bystander. The first type is personal costs for helping which are negative and undesirable consequences on the benefactor who step in and offer help (e.g. physical harm, embarrassment, threat to personal freedom). The other type involves costs for the victim not receiving help (e.g. guilt, fear of negative evaluation by others, and empathic costs as a result of the awareness of victim's continued suffering). Rewards, on the other hand, refer to anticipated positive outcomes of helping (e.g. monetary compensation, social praise). Overall, the basic assumption of this model is derived from the economic perspective of human behavior—people tend to behave in ways that minimize costs and maximize rewards (Dovidio, 1984). From this viewpoint, individuals are relatively rational and self-interested when making decisions. In an emergency, bystanders are motivated to reduce aversive arousal caused by witnessing a victim in distress. Then they analyze the situation, assess the potential costs and rewards associated with alternative behavioral reactions, and subsequently, make a final decision that will incur as few net costs (costs minus rewards) as possible.

As the arousal: cost-reward model suggests, the severity of the situation and the emotional involvement with the victim are critical in influencing the extent to which observation of an emergency arouses a bystander (Dovidio, 1984). In line with this proposition, cyberbystanders are likely to experience unpleasant arousal when witnessing a victim being verbally abused by someone else. The degree of arousal in response to a victim's plight is a function of situational severity, with greater severity producing higher levels of arousal (Van Noorden et al., 2016). Besides, the relationship with a victim may also affect cyberbystanders' levels of arousal and their cost-reward assessments, which further determines the ways they respond to the situation. DeSmet et al. (2014) found that cyberbystanders expressed more willingness to defend the victim who was a friend even though they clearly recognized the potential risk of intervention because they thought the risk of being isolated as a result of not defending a friend as more threatening. Therefore, the following sections focus on how situational severity and interpersonal relationship in terms of similarity dimension influence cyberbystanders' likelihood of intervention by unpacking the psychological mechanisms underlying this process.

Severity of cyberbullying

Cyberbystanders tend to engage in self-assessment process through which they evaluate the severity of the situation before taking actions (Gahagan et al., 2016). Interpretation of the situation as an emergency is an essential step to determine whether bystanders finally intervene on behalf of the victim (Latané and Darley, 1970). Prior research indicates that a highly dangerous emergency is recognized faster and less ambiguously than a less dangerous situation (Fischer et al., 2006). Unambiguous emergency involving severe consequences to the victim promotes bystanders to engage in helping behaviors than does an ambiguous situation (Clark and Word, 1972). Consistent with this tendency in offline scenario, Obermaier et al. (2016) found that exposure to a highly severe cyberbullying incident led cyberbystanders to perceive the situation as more urgent and feel more responsible for helping the victim, which consequently boosted their intention to intervene.

Although more severe emergencies produce greater likelihood of intervention (e.g. Fischer et al., 2011), bystanders tend to recognize a high-severity situation as more of a problem and less safe to intervene in than a low-severity situation (Bennett and Banyard, 2016). In this sense, situational severity may influence the ways in which cyberbystanders manage the visibility of intervening behaviors for the purpose of minimizing the potential costs for helping. Compared to private intervention, publicly stepping in the situation and offering help to the victim may result in more negative consequences. For example, cyberbystanders are concerned about being retaliated and becoming the next victim as a result of publicly defending the victim in front of the perpetrator and other bystanders (DeSmet et al., 2014). From a qualitative case study, Thornberg (2007) found that bystanders often thought that helping a victim in front of others can not only make that victim even more embarrassed but also make themselves being teased or negatively judged by others. Besides, several pieces of qualitative research indicate that cyberbystanders prefer to respond to cyberbullying in private rather than in public as they consider private intervention as a safer way to not further inflame the situation (DeSmet et al., 2014; Patterson et al., 2017). These findings suggest that the costs associated with private intervention are relatively lower than the costs for public intervention. When confronted with a highly severe cyberbullying, it is plausible to infer that cyberbystanders tend to privately, instead of publicly, cope with the situation. This leads to the following hypothesis:

H1. Cyberbystanders who witness a high-severity cyberbullying incident will be (a) less likely to engage in public intervention, and (b) more likely to engage in private intervention than those who witness a low-severity incident.

Empathic response to cyberbullying

According to the arousal: cost-reward model, empathy is a primary motivational construct underlying helping behavior (Piliavin et al., 1981). In a broad sense, empathy refers to one's reactions to the observed experiences of another (Davis, 1983). It describes the ability to vicariously experience emotions and sensations of others (Clark, 1980). Psychological research has indicated that empathy is a multifaceted concept that has two dimensions, namely cognitive empathy and affective empathy (Gini, 2006). Cognitive empathy is the ability to understand another's emotions and take the perspective of that person, and affective empathy is the vicarious emotional responses stemming from another's emotional state and condition (Davis, 1983). Both dimensions are positively associated with cyberbystanders' prosocial behaviors (e.g. defending victim), and negatively related to their anti-social behaviors (e.g. forwarding a mockery message) in cyberbullying (Van Cleemput et al., 2014). This study focuses on the affective dimension of empathy by examining how cyberbystanders' emotional responses from connecting the victim affectively may influence their willingness to intervene in cyberbullying.

The extent to which bystanders are empathically aroused depends on the severity of the situation (Dovidio, 1984). Previous studies have provided empirical evidence for a direct relationship between situational severity and observers' arousal. For example,

Geer and Jarnecky (1973) conducted a laboratory experiment in which participants witnessed victims who were apparently receiving electric shocks of varying levels of intensity. Participants observing higher intensity shocks were more aroused, reflected as faster reaction time and higher numbers of spontaneous skin conductance, than those observing lower intensity shocks. On this basis, it seems plausible that bystanders infer greater pain victims suffered as the severity of an emergency increases. The more severe the situation, the more empathic arousal bystanders experience (Van Noorden et al., 2016).

One may argue that, due to physical isolation and reduced nonverbal cues in the computer-mediated context, the inability to directly see the actual reactions of cyber victims (e.g. painful facial expression) may inhibit cyberbystanders' empathic arousal. However, this is not necessarily the case as several pieces of empirical evidence suggest that cyberbystanders are not absolutely apathetic to victims of cyberbullying, but rather, that their activation of empathy appears to depend on situational factors relating to cyberbullying incidents (e.g. Machackova et al., 2016). Given that relatively severe emergencies capture the attention of bystanders and reduce ambiguity of the situation (Dovidio, 1984), cyberbystanders who witness a victim being harshly insulted online are more likely to unambiguously recognize the serious consequences of the incident and feel a greater sense of empathic arousal than those who witness a mildly severe situation.

Furthermore, empathic response to witnessing the victim in distress leads bystanders to not only show concern for the needy victim but also feel personal discomfort and anxiety (Batson and Coke, 1981). To mitigate such unpleasant feelings, bystanders are motivated to take certain actions. "The most direct way to end one's vicarious distress is typically to end the victim's troubles" (Piliavin et al., 1981: 49). Bystanders who feel empathy for a victim tend to act in a way beneficial to that victim. Past studies have demonstrated that empathy in response to a victim's plight is a potent motivator of bystander prosocial behavior in offline bullying episodes (e.g. Batson and Shaw, 1991; Gini et al., 2007). In accordance with this pattern, Van Cleemput et al. (2014) showed that cyberbystanders with higher levels of empathic response were more likely to help victims of cyberbullying, while those with lower levels of empathic response were more inclined to remain outsiders, or even join in the situation. Following this, empathy may serve as an underlying mechanism that mediates the indirect effects of cyberbullying severity on willingness to intervene publicly and privately. Therefore, the following hypothesis is raised:

H2. Cyberbystanders who witness a high-severity cyberbullying incident will (a) experience greater levels of empathy, and subsequently show (b) lesser likelihood of public intervention and (c) greater likelihood of private intervention than those who witness a low-severity incident.

Interpersonal similarity between bystander and victim

Another factor influencing cyberbystanders' willingness to intervene is their similarity to the victim. Interpersonal similarity is a form of social distance—the more similar bystanders perceive the victim is with themselves, the shorter social distance they would have with the victim (Liviatan et al., 2008). Victims of cyberbullying on social media can be within or outside of cyberbystanders' own social circle. Although social networks

maintained on social media (e.g. Facebook) are largely originated from the relationship initiated offline, it is not surprising to see that social media are often used to form and connect with loose social ties (“weak ties”), such as “friends-of-friends” or even strangers (Ellison et al., 2007). Social network sites afford users to form impressions of others through the information others disclosed online (e.g. profile descriptions, photographs, content posted and shared). Sprecher (2014) found that individuals who received bogus information that the interacting other was more similar to themselves experienced more closeness and liking toward the other. Given this, levels of similarity to a person exert an impact on cyberbystanders’ affective reactions to that person, which may influence their reactive behaviors when witnessing that person being bullied online.

While research on offline prosocial behavior has revealed the tendency that bystanders are more likely to help others who are more similar (e.g. Batson and Shaw, 1991; Gray et al., 1991), few empirical studies have examined the role of similarity in online emergency situation, such as cyberbullying. In the light of the arousal: cost-reward model, similarity between bystander and victim can facilitate the formation of *we-ness*, referring to “a sense of connectedness or a categorization of another person as a member of one’s own group” (Dovidio et al., 1991: 102). In other words, when an individual perceives a greater similarity to others, whether in beliefs, attitudes, personality, or other characteristics, that individual tends to categorize others as members of in-group. Such categorization increases the feelings of closeness and perceptions of personal responsibility for the welfare of others. That, in turn, elicits greater physiological arousal and empathy in response to the distress of more similar others (Batson et al., 1995; Krebs, 1975) as well as higher costs for not helping (e.g. feelings of guilt). Moreover, several research shows that similarity facilitates empathy for a target person such that individuals experience more empathy for similar others than for dissimilar others (Brown et al., 2006; Cialdini et al., 1997). When combined with situational severity, similarity with a victim may activate stronger feelings of empathy for a more similar victim involved in a high-severity cyberbullying than in a low-severity cyberbullying. Such empathic response may, in turn, affect cyberbystanders’ willingness to intervene. This possibility is examined in the following hypothesis:

H3. The indirect effects of high-severity cyberbullying (in contrast to low-severity) on (a) public or (b) private intervention through increased levels of empathy will be more pronounced when cyberbystanders have a higher level of similarity with the victim.

A proposed model encompassing the hypotheses is illustrated as Figure 1.

Methods

Participants

One hundred and forty-nine participants were recruited from Amazon Mechanical Turk. After excluding participants who failed the manipulation check ($N=9$) and those who were not Facebook users ($N=8$), the final sample size was 132 (74 males, 58 females;

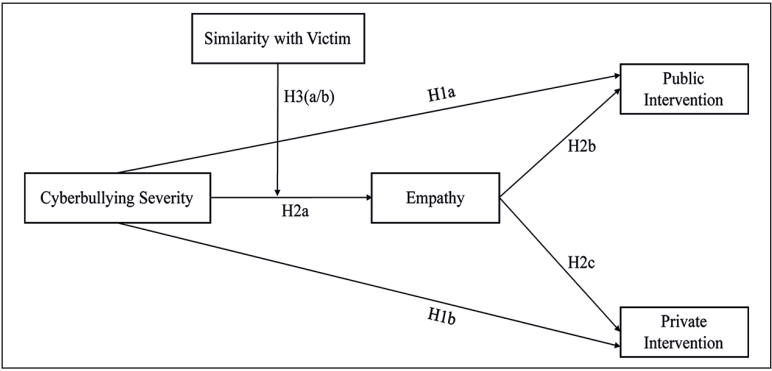


Figure 1. Proposed research model.

$M_{age}=37.75$, $SD_{age}=11.36$). The majority of participants identified themselves as Caucasians (79.5%). In terms of educational level, 51.5% of participants held a bachelor’s degree or above.

Study design

An online experiment with a 2 (cyberbullying severity: high vs low) $\times$ 2 (similarity with the victim: high vs low) between-subjects design was conducted. Participants were randomly assigned to one of the four conditions.

Stimuli

The stimuli were created within the template of Facebook. It depicted a fictitious cyberbullying incident on Facebook where Alex (victim) posted a status update about personal issue and Taylor (bully) left a comment attacking Alex. These unisex names were created for the study in order to minimize the potential impact of gender on participants’ reactions. Moreover, additional information about victim and bully (e.g. last name, profile picture) were blurred out.

To enhance external validity and avoid case-category confounding issue as a result of using single stimulus (Jackson, 1992), this study created two fictitious scenarios of cyberbullying. One scenario illustrated that Alex’s team lost the game but they learned a lot from this experience (“game” scenario, hereafter), and another was about that Alex wanted to lose weight but felt too tired to work out (“weight” scenario, hereafter). The length and tone of Alex’s posts was kept consistent across different scenarios.

To determine cyberbullying messages to be embedded in experimental stimuli, a pilot test was conducted among 55 participants. Given that cyberbullying often marks harsh insults on social networking sites (Livingstone et al., 2013), 12 insulting messages directed at Alex were created for each scenario. Participants were randomly assigned to one of the two scenarios, and asked to evaluate 12 messages presented as a reply with regard to several dimensions adapted from previous works (e.g. Brody and Vangelisti,

2016; Kazerooni et al., 2018). These dimensions included perceived presence of cyberbullying, perceived intention to harm the recipient, perceived severity, perceived valence, and perceived believability.

As a result of the pilot test, four messages, two for each scenario, were selected for the experimental stimuli representing the high and low severity. For the “game” scenario, the reply with high severity was “You seriously deserve to die,” while the reply in the low-severity condition was “You’d better not play if you do not know how to play!” For the “weight” scenario, highly severe reply stated, “You looks disgusting, fat bastard!” and the reply of low severity was “You are really good at finding excuse for yourself.” Furthermore, the independent sample *t*-tests confirmed that there were no significant differences between two messages of varying severity across different scenarios in terms of the five dimensions mentioned above. The examples of the stimuli are shown in Supplemental Appendix A.

Procedure

After agreeing to participate in the study by signing the consent form, participants were first presented with a cover story indicating that this study intends to examine users’ experiences with a similarity-test feature designed for Facebook in order to help Facebook users find friends of similar characteristics and interests. Participants were then informed that another Facebook user—Alex—was also invited to participate in the similarity test, and the purpose of this test was to evaluate the extent to which they have in common with Alex. After this, participants responded to a self-descriptive form by answering some questions about themselves on dichotomous items. The questions included “Which do you prefer?” (e.g. coffee vs tea; city vs countryside) and “Which best describes you?” (e.g. introvert vs extrovert; organized vs easygoing).

Upon completion of the self-descriptive form, participants were instructed to wait a few seconds for the system to calculate the similarity between Alex and themselves. Participants in the high similarity condition were informed that they have 90% similarities with Alex, whereas those in the low similarity condition were told that they have 10% similarities with Alex. Our approach of similarity manipulation was adapted from previous works (e.g. Sprecher, 2014; Wei and Liu, 2020).

Next, participants were randomly assigned to read a screenshot that consisted of a fictitious Facebook post where Alex described a personal issue with a reply by Taylor. Immediately after reading, participants completed a questionnaire including manipulation check items, dependent measures, and demographic information. Finally, participants were debriefed about the nature of the study.

Measures

Empathy. Five items adapted from Davis (1983) were used to measure emotional responses from the discomfort and anxiety of the victim. Participants were asked to indicate the extent to which they felt “worried,” “uneasy,” “distressed,” “troubled,” and “upset” when reading Taylor’s reply, on a 5-point Likert-type scale (1 = *not at all*, 5 = *extremely*) ($M = 3.08$, standard deviation [SD] = 1.33, Cronbach’s $\alpha = .95$).

Public intervention. Four items adapted from Bastiaensens et al. (2014) were employed to assess participants' intention to publicly engage in intervention. Participants were asked to indicate the extent to which they would like to "Publicly tell Alex you think that Taylor's reply is not OK in a comment on Facebook (other people on Facebook can see your comment)," "Publicly comfort Alex in a Facebook comment," "Publicly give Alex advice on what to do in a Facebook comment," and "Publicly post a Facebook comment telling Taylor to stop being mean" on a 7-point Likert-type scale (1 = *extremely unlikely*, 7 = *extremely likely*) ($M=4.15$, $SD=1.80$, Cronbach's $\alpha=.90$).

Private intervention. Items assessing the intention to privately intervene in the cyberbullying incident were slightly modified from the measures of public intervention. Four items included "Privately tell Alex you think that Taylor's reply is not OK via SMS, chat or email (only between you and Alex, other people cannot see it)," "Privately comfort Alex via SMS, chat or email," "Privately give Alex advice on what to do via SMS, chat or email," and "Privately tell Taylor to stop being mean via SMS, chat or email" (1 = *extremely unlikely*, 7 = *extremely likely*) ($M=4.37$, $SD=1.78$, Cronbach's $\alpha=.89$).

Control variables. Several variables, including dispositional empathy (Davis, 1983), personal experience with cyberbullying (i.e. frequency of being cyberbullied by others, frequency of cyberbullying others, frequency of witnessing a cyberbullying incident), hours per day on Facebook, number of posts updated on Facebook daily, and demographic information (i.e. gender, age, race, religion, educational level), were considered as control variables. Among these variables, gender, dispositional empathy, and hours on Facebook per day significantly affected the findings reported herein, so they were included as covariates in the analyses.

Results

Manipulation check

Participants were asked to recall the similarity score between Alex and themselves calculated by algorithm, and selected the answer from "10%," "20%," "80%," and "90%." Those who failed to choose the right answer were excluded from final analysis ($N=9$). Moreover, two items adapted from Sprecher (2014) were used to assess participants' perceived similarity with Alex. Sample items included: "How much do you think you have in common with Alex?" and "How similar do you think you and Alex are likely to be?" Participants responded to the items on a 5-point Likert-type scale (1 = *not at all*, 5 = *a great deal*) ($M=2.78$, $SD=1.08$, Cronbach's $\alpha=.96$). The results of an independent sample *t*-test showed that participants in the high similarity condition ($M=3.33$, $SD=1.03$) perceived significantly greater similarity with the victim than did participants in the low similarity condition ($M=2.18$, $SD=0.77$), $t(123.59)=7.30$, $p<.001$. Thus, the manipulation of similarity was considered as successful.

Regarding the varying severities of cyberbullying, participants were asked to indicate the degree to which they think the reply by Taylor was "severe," "threatening," "dangerous," "alarming," and "hurtful" on a 5-point Likert-type scale (1 = *not at all*, 5 = *extremely*)

Table 1. Number of participants assigned to each condition.

	Low similarity	High similarity	Total
Low severity	30	35	65
High severity	34	33	67
Total	64	68	132

Table 2. Bivariate correlations among measured variables.

	1	2	3	4	5	6
1. Empathy	—					
2. Public intervention	.53**	—				
3. Private intervention	.53**	.67**	—			
4. Dispositional empathy	.42**	.44**	.47**	—		
5. Gender	.20*	.09	.06	.14	—	
6. Time spent on Facebook per day	.26**	.39**	.38**	.17	.07	—

* $p < .05$; ** $p < .01$.

(Obermaier et al., 2016) ($M = 3.30$, $SD = 1.27$, Cronbach’s $\alpha = .91$). The analysis revealed a significant difference between high-severity ($M = 3.99$, $SD = 1.10$) and low-severity condition ($M = 2.58$, $SD = 1.04$), $t(130) = 7.52$, $p < .001$, confirming the manipulation of severity was successful.

Preliminary analysis

PROCESS macro (Hayes, 2017, Model 58) was employed to examine whether the different scenarios would influence the effects of independent variable (i.e. cyberbullying severity) on mediator (i.e. empathy) and the effects of mediator on dependent variables (i.e. public and private intervention). The results indicated that there was no significant interaction involving the scenario, all $ps > .05$. Thus, the data were collapsed across the two scenarios for the following analyses (see Table 1). The intercorrelations among key variables are presented in Table 2.

Hypothesis testing

Two-way analyses of variance (ANOVAs) were conducted to test H1 and H2a. Participants in high-severity condition ($M = 3.50$, standard error [SE] = .13) experienced greater levels of empathy than those in low-severity condition ($M = 2.63$, $SE = .14$), $F(1, 125) = 19.93$, $p < .05$, partial $\eta^2 = .14$, lending support to H2a. The high-severity condition ($M = 4.43$, $SE = .19$) elicited a greater intention of public intervention than the low-severity condition ($M = 3.86$, $SE = .19$), $F(1, 125) = 10.28$, $p < .05$, partial $\eta^2 = .04$. Contrary to the prediction, participants displayed more willingness to privately intervene when they witnessed a high-severity incident ($M = 4.76$, $SE = .18$) as opposed to a

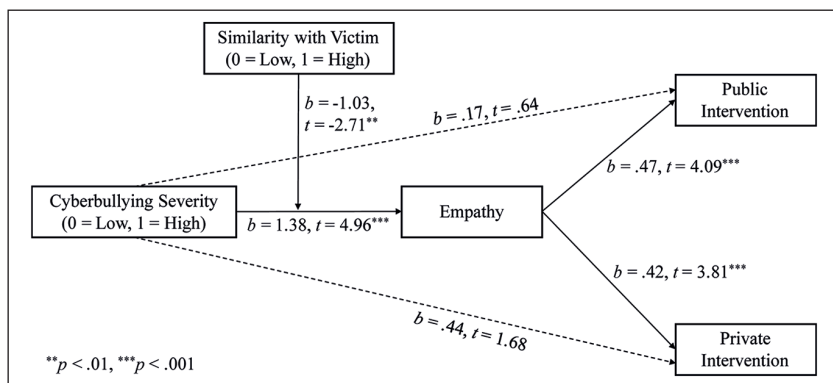


Figure 2. Conditional indirect effects of cyberbullying severity on public and private intervention via empathy.

low-severity one ($M = 3.94, SE = .18$), $F(1, 125) = 21.10, p < .01$, partial $\eta^2 = .08$. Therefore, H1b was supported, but H1a was rejected.

To test the indirect effects of empathy proposed in H2b-c, Model 4 of the PROCESS macro (Hayes, 2017) with 10,000 bootstrap samples and 95% bias-corrected confidence intervals (CIs) was used. Degree of severity was dummy coded (1 = high severity, 0 = low severity). Gender, dispositional empathy, and hours spent on Facebook per day were included as covariates in the analyses. Results indicated that empathy significantly mediated the effect of cyberbullying severity on public intervention (index = .39, $SE = .16$, 95% CI = [.14, .75]) and on private intervention (index = .36, $SE = .14$, 95% CI = [.13, .68]). Participants who witnessed the more severe cyberbullying experienced greater levels of empathy than those who witnessed the less severe cyberbullying ($b = .85, t = 4.30, p < .001$). Such empathy, in turn, promoted participants' likelihood of public intervention ($b = .47, t = 4.09, p < .001$), $F(4, 127) = 15.31, p < .001, R^2 = .33$, and private intervention ($b = .42, t = 3.81, p < .001$), $F(4, 127) = 17.89, p < .001, R^2 = .36$. Thus, H2c was supported, yet H2b was not confirmed. Among control variables, dispositional empathy was significantly and positively associated with public intervention ($b = .43, t = 3.36, p = .001$) as well as private intervention ($b = .50, t = 3.99, p < .001$). The time spent on Facebook per day also positively predicted cyberbystanders' likelihood of intervening (for public intervention: $b = .39, t = 3.53, p < .001$; for private intervention: $b = .34, t = 3.18, p < .01$).

Concerning the moderating effects of the similarity with the victim, moderated mediation analyses were performed by using Model 7 (see Figure 2). The indirect effect of cyberbullying severity on public intervention through empathy were significantly moderated by the similarity with the victim, but this only held when the similarity level was low ($b = .64, SE = .24$, 95% CI = [.24, 1.17]), not when it was high ($b = .16, SE = .15$, 95% CI = [-.09, .50]), $F(5, 126) = 17.11, p < .001, R^2 = .40$. Contrary to our hypothesis, participants exhibited greater empathy for the dissimilar victim in the high-severity condition ($M = 3.84, SE = .19$) than in the low-severity condition ($M = 2.46, SE = .20$), $F(1, 125) = 7.36, p < .01$ (see Figure 3). Such increased empathy, in turn, led to higher intention to publicly

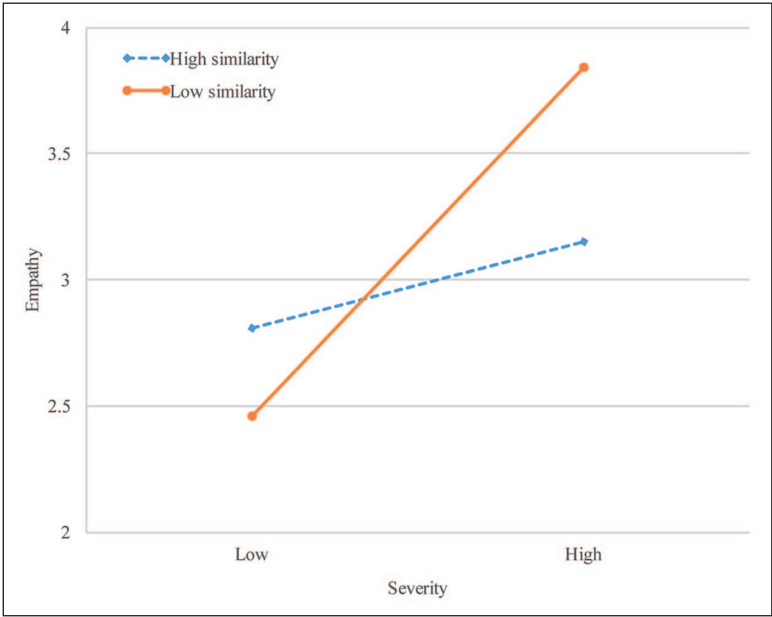


Figure 3. Interaction between cyberbullying severity and interpersonal similarity on empathy.

intervene ($b=.47, t=4.09, p<.001$). In addition, the index of moderated mediation showed that the similarity with the victim acted as a significant moderator (index= $-.48, SE=.23, 95\% CI=[-1.00, -.10]$), but in the opposite direction of H3a.

The degree of similarity between victim and cyberbystander also played a significant role in moderating the indirect effect of cyberbullying severity on private intervention via the perceived empathy (index= $-.43, SE=.21, 95\% CI=[-.90, -.09]$), but in the opposite direction of H3b. That is, this conditional indirect effect was only statistically significant when the similarity level was low ($b=.58, SE=.22, 95\% CI=[.21, 1.07]$), not when it was high ($b=.15, SE=.13, 95\% CI=[-.08, .46]$), $F(5, 126)=18.75, p<.001, R^2=.43$. Contrary to the proposed hypothesis, high-severity cyberbullying involving a dissimilar victim elicited higher levels of empathy than the low-severity situation involving the same victim, which subsequently promoted likelihood of private intervention ($b=.42, t=3.81, p<.001$).

Discussion

The present research investigated how cyberbullying severity and interpersonal similarity influenced cyberbystander likelihood of intervening in cyberbullying on social media. Consistent with the predictions derived from the arousal: cost-reward model (Dovidio et al., 1991; Piliavin et al., 1981), cyberbystanders’ empathic responsiveness depended on the varying degrees of cyberbullying severity. The more empathy bystanders felt, the more likely they were to publicly and privately intervene. In addition, the indirect effect of cyberbullying severity on willingness to intervene through empathy was more

pronounced when the interpersonal similarity between the victim and the cyberbystander was low, rather than it was high.

Unlike most existing research on cyberbystander that has relied on Latané and Darley's (1970) bystander intervention model, this study provides a novel theoretical perspective to advance current understanding of cyberbystander behavior on social media by applying the use of the arousal: cost-reward model in offline settings to online situations. Our findings explicate the mediating effect of empathic arousal as the principal mechanism underlying intervention behavior. More intriguingly, empathy fully mediated the indirect effect of incident severity on public and private forms of intervention, which might provide an explanation concerning apathetic reactions of onlookers. Observation of cyberbullying did not directly lead cyberbystanders to step in and offer help; instead, their intervention was primarily motivated by a desire to reduce unpleasant arousal with the distress of the victim. This provides supportive evidence for the arousal: cost-reward model in an online emergent situation (i.e. cyberbullying) by showing that empathic arousal in response to the victimization of others is the central motivational construct underlying helping behavior. When empathy was evoked by witnessing a victim being bullied online, cyberbystanders relied primarily on their affective state to engage in impulsive, noncalculative intervention, without apparent regard for the costs associated with different forms of intervention (e.g. being retaliated by perpetrator). That might explain why cyberbystanders experiencing empathy were willing to intervene both privately and publicly.

Another important finding of this study is that cyberbystander's similarity to the victim moderated the indirect effect of cyberbullying severity on intervening intention via empathy. Compared to those who witnessed a dissimilar victim in a low-severity cyberbullying, cyberbystanders who witnessed a dissimilar victim involved in a high-severity cyberbullying experienced higher levels of empathy and thus expressed more willingness to intervene publicly and privately. When the victim was similar to cyberbystanders, however, their likelihood of engaging in intervention did not vary significantly between different severity conditions. This unanticipated tendency could be possibly explained by the victimization perspective (Wortman and Dunkel-Schetter, 1979). According to this perspective, individuals sometimes tend to exhibit unsupportive behaviors (e.g. avoidance) in a victimizing event because of the feelings of negative emotions (e.g. fear, anxiety, insecurity, vulnerability). For instance, Artiss and Levine (1973) found that doctors felt anxious about encountering dying patients and avoided interacting with them. Similarly, another experiment shows that participants who met with a cancer patient for the first time were more likely to experience uncomfortable feelings and display avoidance behavior than those who met with a healthy person (Silver et al., 1990). Research guided by self-categorization approach (Turner et al., 1987) indicates that the level of shared group membership or similarity between self and other affects the degree to which individuals perceive what happens to others may be personally relevant. When the victim of an emergency is described as an in-group rather than an out-group member, individuals are more likely to feel personally threatening and infer that the same fate might quite easily befall them (Stapel et al., 1994). Likewise, if witnessing a cyberbullying incident involving more severe consequences to a victim, perceiving oneself as similar to that victim may exacerbate the feelings of anxiety and vulnerability by making cyberbystanders feel that the

victimization might befall themselves as well, which possibly inhibits experience of empathy and subsequent motivation to intervene.

Although this study focused on the role of situational empathy aroused from witnessing distress and sufferings of the victim, results also revealed that dispositional empathy was a robust predictor of cyberbystander intervention. This is consistent with the findings shown in previous research that dispositional empathy of cyberbystanders is positively associated with helping behavior (e.g. Freis and Gurung, 2013). It provides practical implications for cyberbullying intervention program by demonstrating the importance of empathy training in converting bystanders to upstanders. For instance, raising awareness of the serious consequences of cyberbullying victimization may be an effective strategy that can help online users to sense and experience victim's sufferings and facilitate their activation of empathy into actual intervening behavior.

Limitations and directions for future research

Several limitations of this study merit further investigation. First, we manipulated interpersonal similarity through bogus trait information with an unacquainted individual. Although this approach seems to produce more realism compared to previous experimental studies simply asking participants to imagine their proximity to the victim of cyberbullying (e.g. Bastiaensens et al., 2014; Brody and Vangelisti, 2016), it still faces the problem concerning ecological validity as participants might realize that they were in a simulated context. Their reactions in the experimental setting might differ from what they would behave in a real online interaction. The trade-off between internal and external validity is a rather inherent challenge in experimental methods, but future researchers should be more cautious about it when designing the experiment to investigate the role of interpersonal relationship in cyberbystander behavior. Besides, similarity with another person can also be assessed from various aspects, such as gender similarity or attitudinal similarity on controversial topics. Past research suggests that experimental manipulation of similarity can produce differences in empathic arousal (e.g. Batson et al., 1995; Krebs, 1975). Future research should test the robustness of the moderating effects of similarity reported in this study by employing other approaches of similarity manipulation.

A second limitation concerns the generalizability of findings in this study. The two fictitious scenarios of cyberbullying used in the experiment were quite similar—both featuring an insulting reply to a victim's self-disclosure of personal experience or feeling via Facebook's status update. The effects observed herein might not generalize to other types of cyberbullying (e.g. spreading false rumors about a victim). Therefore, future research should validate our findings by using varying forms and contexts of cyberbullying.

Finally, the primary focus of this study is on affective motivational construct of the arousal: cost-reward model (i.e. empathic arousal). Future empirical research should take cost-reward construct into consideration to further explore the cognitive processes through which cyberbystanders appraise potential costs and rewards associated with their courses of action in cyberbullying. Although the affective and cognitive constructs are conceptually distinct, they are functionally interdependent (Dovidio, 1984). It would be worthwhile to investigate the interplay between these two components, such as how empathic response to cyberbullying produces variations in assessed costs and rewards,

and how cost-reward assessment, in turn, affects the extent to which cyberbystander are empathically aroused.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Sai Wang  <https://orcid.org/0000-0001-8562-9470>

Supplemental material

Supplemental material for this article is available online.

References

- Anderson J, Bresnahan M and Musatics C (2014) Combating weight-based cyberbullying on Facebook with the dissenter effect. *Cyberpsychology, Behavior, and Social Networking* 17(5): 281–286.
- Artiss KL and Levine AS (1973) Doctor-patient relation in severe illness: a seminar for oncology fellows. *The New England Journal of Medicine* 288(23): 1210–1214.
- Bastiaensens S, Vandebosch H, Poels K, et al. (2014) Cyberbullying on social network sites. An experimental study into bystanders' behavioural intentions to help the victim or reinforce the bully. *Computers in Human Behavior* 31: 259–271.
- Batson CD and Coke JS (1981) Empathy: a source of altruistic motivation for helping? In: Rushton JP and Sorrentino RM (eds) *Altruism and Helping Behavior: Social, Personality, and Developmental Perspectives*. Hillsdale, NJ: Lawrence Erlbaum, pp. 167–187.
- Batson CD and Shaw LL (1991) Evidence for altruism: toward a pluralism of prosocial motives. *Psychological Inquiry* 2(2): 107–122.
- Batson CD, Turk CL, Shaw LL, et al. (1995) Information function of empathic emotion: learning that we value the other's welfare. *Journal of Personality and Social Psychology* 68(2): 300–313.
- Bennett S and Banyard VL (2016) Do friends really help friends? The effect of relational factors and perceived severity on bystander perception of sexual violence. *Psychology of Violence* 6(1): 64–72.
- Brody N and Vangelisti AL (2016) Bystander intervention in cyberbullying. *Communication Monographs* 83(1): 94–119.
- Brown LM, Bradley MM and Lang PJ (2006) Affective reactions to pictures of ingroup and out-group members. *Biological Psychology* 71(3): 303–311.
- Chen L, Ho SS and Lwin MO (2017) A meta-analysis of factors predicting cyberbullying perpetration and victimization: from the social cognitive and media effects approach. *New Media & Society* 19(8): 1194–1213.
- Cialdini RB, Brown SL, Lewis BP, et al. (1997) Reinterpreting the empathy-altruism relationship: when one into one equals oneness. *Journal of Personality and Social Psychology* 73(3): 481–494.
- Clark KB (1980) Empathy: a neglected topic in psychological research. *American Psychologist* 35(2): 187–190.
- Clark RD and Word LE (1972) Why don't bystanders help? Because of ambiguity? *Journal of Personality and Social Psychology* 24(3): 392–400.

- Davis MH (1983) Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology* 44(1): 113–126.
- DeSmet A, Veldeman C, Poels K, et al. (2014) Determinants of self-reported bystander behavior in cyberbullying incidents amongst adolescents. *Cyberpsychology, Behavior, and Social Networking* 17(4): 207–215.
- Dillon KP and Bushman BJ (2015) Unresponsive or un-noticed? Cyberbystander intervention in an experimental cyberbullying context. *Computers in Human Behavior* 45: 144–150.
- Dovidio JF (1984) Helping behavior and altruism: an empirical and conceptual overview. In: Berkowitz L (ed.) *Advances in Experimental Social Psychology*, vol. 17. New York: Academic Press, pp. 361–427.
- Dovidio JF, Piliavin JA, Gaertner S, et al. (1991) The arousal: cost-reward model and the process of intervention: a review of the evidence. In: Clark MS (ed.) *Review of Personality and Social Psychology*. Thousand Oaks, CA: SAGE, pp. 86–118.
- Ellison NB, Steinfield C and Lampe C (2007) The benefits of Facebook “friends”: social capital and college students’ use of online social network sites. *Journal of Computer-Mediated Communication* 12(4): 1143–1168.
- Fischer P, Greitemeyer T, Pollozek F, et al. (2006) The unresponsive bystander: are bystanders more responsive in dangerous emergencies? *European Journal of Social Psychology* 36(2): 267–278.
- Fischer P, Krueger JJ, Greitemeyer T, et al. (2011) The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin* 137(4): 517–537.
- Freis SD and Gurung RA (2013) A Facebook analysis of helping behavior in online bullying. *Psychology of Popular Media Culture* 2(1): 11–19.
- Gahagan K, Vaterlaus JM and Frost LR (2016) College student cyberbullying on social networking sites: conceptualization, prevalence, and perceived bystander responsibility. *Computers in Human Behavior* 55: 1097–1105.
- Geer JH and Jarnecky L (1973) The effect of being responsible for reducing another’s pain on subjects’ response and arousal. *Journal of Personality and Social Psychology* 26(2): 232–237.
- Gini G (2006) Social cognition and moral cognition in bullying: what’s wrong? *Aggressive Behavior* 32(6): 528–539.
- Gini G, Albiero P, Benelli B, et al. (2007) Does empathy predict adolescents’ bullying and defending behavior? *Aggressive Behavior* 33(5): 467–476.
- Gray C, Russell P and Blockley S (1991) The effects upon helping behaviour of wearing pro-gay identification. *British Journal of Social Psychology* 30(2): 171–178.
- Hayes AF (2017) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. 2nd ed. New York: Guilford Press.
- Huang YY and Chou C (2010) An analysis of multiple factors of cyberbullying among junior high school students in Taiwan. *Computers in Human Behavior* 26(6): 1581–1590.
- Jackson S (1992) *Message Effects Research: Principles of Design and Analysis*. New York: Guilford Press.
- Kazerooni F, Taylor SH, Bazarova NN, et al. (2018) Cyberbullying bystander intervention: the number of offenders and retweeting predict likelihood of helping a cyberbullying victim. *Journal of Computer-Mediated Communication* 23(3): 146–162.
- Kowalski RM and Limber SP (2007) Electronic bullying among middle school students. *Journal of Adolescent Health* 41(6): 22–30.
- Krebs D (1975) Empathy and altruism. *Journal of Personality and Social Psychology* 32(6): 1134–1146.
- Latané B and Darley JM (1970) *The Unresponsive Bystander: Why Doesn’t He Help?* Englewood Cliffs, NJ: Prentice Hall.

- Leary MR and Kowalski RM (1990) Impression management: a literature review and two-component model. *Psychological Bulletin* 107(1): 34–47.
- Liviatan I, Trope Y and Liberman N (2008) Interpersonal similarity as a social distance dimension: implications for perception of others' actions. *Journal of Experimental Social Psychology* 44(5): 1256–1269.
- Livingstone S, Ólafsson K and Staksrud E (2013) Risky social networking practices among “underage” users: lessons for evidence-based policy. *Journal of Computer-Mediated Communication* 18(3): 303–320.
- Lynn R (1966) *Attention, Arousal and the Orientation Reaction*. Oxford: Pergamon Press.
- Machackova H, Dedkova L, Sevcikova A, et al. (2016) Empathic responses by cyberbystanders: the importance of proximity. *Journal of Youth Studies* 19(6): 793–804.
- O'Connell P, Pepler D and Craig W (1999) Peer involvement in bullying: insights and challenges for intervention. *Journal of Adolescence* 22(4): 437–452.
- Obermaier M, Fawzi N and Koch T (2016) Bystanding or standing by? How the number of bystanders affects the intention to intervene in cyberbullying. *New Media & Society* 18(8): 1491–1507.
- Patchin JW and Hinduja S (2012) *Cyberbullying Prevention and Response: Expert Perspectives*. New York: Routledge.
- Patterson LJ, Allan A and Cross D (2017) Adolescent perceptions of bystanders' responses to cyberbullying. *New Media & Society* 19(3): 366–383.
- Piliavin JA, Dovidio JF, Gaertner SL, et al. (1981) *Emergency Intervention*. New York: Academic Press.
- Salmivalli C (2010) Bullying and the peer group: a review. *Aggression and Violent Behavior* 15(2): 112–120.
- Silver RC, Wortman CB and Crofton C (1990) The role of coping in support provision: the self-presentational dilemma of victims of life crises. In: Sarason IG, Sarason BR and Pierce GR (eds) *Social Support: An Interactional View*. New York: John Wiley, pp. 397–426.
- Smith PK, Mahdavi J, Carvalho M, et al. (2008) Cyberbullying: its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry* 49(4): 376–385.
- Sprecher S (2014) Effects of actual (manipulated) and perceived similarity on liking in get-acquainted interactions: the role of communication. *Communication Monographs* 81(1): 4–27.
- Stapel DA, Reicher SD and Spears R (1994) Social identity, availability and the perception of risk. *Social Cognition* 12(1): 1–17.
- Thornberg R (2007) A classmate in distress: schoolchildren as bystanders and their reasons for how they act. *Social Psychology of Education* 10(1): 5–28.
- Tokunaga RS (2010) Following you home from school: a critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior* 26(3): 277–287.
- Turner JC, Hogg MA, Oakes PJ, et al. (1987) *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford: Basil Blackwell.
- Van Cleemput K, Vandebosch H and Pabian S (2014) Personal characteristics and contextual factors that determine “helping,” “joining in,” and “doing nothing” when witnessing cyberbullying. *Aggressive Behavior* 40(5): 383–396.
- Van Noorden TH, Bukowski WM, Haselager GJ, et al. (2016) Disentangling the frequency and severity of bullying and victimization in the association with empathy. *Social Development* 25(1): 176–192.
- Wei L and Liu B (2020) Reactions to others' misfortune on social media: effects of homophily and publicness on schadenfreude, empathy, and perceived deservingness. *Computers in Human Behavior* 102: 1–13.

- Wortman CB and Dunkel-Schetter C (1979) Interpersonal relationships and cancer: a theoretical analysis. *Journal of Social Issues* 35(1): 120–155.
- Ybarra ML and Mitchell KJ (2004) Youth engaging in online harassment: associations with caregiver-child relationships, Internet use, and personal characteristics. *Journal of Adolescence* 27(3): 319–336.

Author biography

Sai Wang (MA, The University of Texas at Austin) is a PhD candidate in the Department of Media and Communication at the City University of Hong Kong, Hong Kong. Her research examines how social interactions via various communication technologies and services influence user psychology and behavior. Her works have been published in the *Journal of Product and Brand Management* and *Journal of Interactive Advertising*.