

Attacks by Anons: A Content Analysis of Aggressive Posts, Victim Responses, and Bystander Interventions on a Social Media Site

Social Media + Society
January-March 2018: 1–14
© The Author(s) 2018
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2056305118762444
journals.sagepub.com/home/sms


Rachel Young¹ , Stephanie Miles¹, and Saleem Alhabash²

Abstract

Cyberbullying is an area of great anxiety related to adolescents' use of social media. Although the affordances of social media sites often allow aggressive online content to be stored and searched, the actual content of aggressive posts has not been explored in great detail. The purpose of this content analysis was to examine discursive strategies used in aggressive posts, responses, and bystander comments on a social media site that is both popular among young adolescents and a known online site of cyberbullying behavior. A total of 993 question–answer dyads were analyzed. In this sample, aggressors are almost exclusively anonymous. Posters of aggressive content demean profile owners on the basis of social status, romantic success, emotional instability, perceived physical attractiveness, and age. Most profile owners attracted a comment from at least one supportive bystander. In general, bystander comments either attacked aggressive posters for their cruelty or their cowardice at being anonymous or supported profile owners by affirming their physical attractiveness or social competence. A power differential between aggressor and victim is a key feature that distinguishes bullying from other social conflicts among adolescents. Results show that, in the absence of physical power, online aggressors use discursive strategies to affirm their dominance over their victims. In turn, victims of online aggression, as well as supportive bystanders, use a variety of methods to attempt to resolve the power differential. Our findings have implications for development of digital citizenship and anti-cyberbullying initiatives that promote effective bystander behavior online.

Keywords

cyberbullying, online aggression, social media, anonymity, bystander

As digital media become increasingly central to adolescent social life, public dialogue about online aggression has also increased. Both digital and traditional forms of bullying are linked to a variety of negative outcomes including poor academic performance, low self-esteem, emotional disorders, and psychosocial problems that may follow adolescents into adulthood (Tokunaga, 2010). Between 15% and 40% of adolescents report having been victims of cyberbullying, and 10% and 20% report having attacked others online (Hinduja & Patchin, 2010; Tokunaga, 2010). The wide range in prevalence estimates reflects variance in the definition and operationalization of cyberbullying (Ybarra, boyd, Korchmaros & Oppenheim, 2012).

Ybarra et al. (2012) suggest that cyberbullying is not a distinct phenomenon but is a type of aggressive interaction that varies according to social presence and other affordances of digital media. Communication online can be asynchronous,

permanent, public, and anonymous, with the potential to reach a widespread audience and draw comments from bystanders both known and unknown to the victim. The phenomenon of cyberbullying has attracted public concern, and social media sites have drawn criticism as media blame the anonymity afforded by aggressors and a lack of site intervention for suicides that occur after documented cyberbullying (e.g., Smith-Spark, 2013). In Festl and Quandt's (2013) study, harassment via social media accounted for more than half of reported cyberbullying.

¹The University of Iowa, USA

²Michigan State University, USA

Corresponding Author:

Rachel Young, School of Journalism and Mass Communication, University of Iowa, 100 Adler Journalism Building, Iowa City, IA 52240, USA.
Email: rachel-young@uiowa.edu



Research into the social function of bullying has found that a motivation for the behavior is to gain status by demonstrating social power (Sutton & Smith, 1999; Veenstra et al., 2007). Because online social networks often reflect the connections forged offline, issues of power and status within social hierarchies may be reflected or played out on social media sites. However, demonstrating power online may be much different than demonstrating power in face-to-face bullying.

While research across disciplines has assessed rates of cyberbullying and proposed prevention strategies, there has been comparatively little focus on the discursive content of online aggression. This study uses content analysis to investigate how attackers, victims, and bystanders negotiate issues of power and social status in cyberbullying interactions online. We examine 993 question–answer dyads from a social media site used predominantly by adolescents to investigate how aggressors use social media affordances and discursive strategies to exhibit social power and, in turn, the strategies victims and bystanders use to respond. Because power is a central construct in defining bullying and now cyberbullying, it is important for researchers to conduct investigations of how power is displayed or constructed in an online environment. Investigating power in cyberbullying discourse can help us understand how words do harm and how victims and bystanders strike back.

Cyberbullying: A Question of Power

Power is a crucial construct in bullying research. Olweus (1978) emphasizes not only that bullies intend to hurt their victims repeatedly over time but also that aggressors perceive themselves to be of greater social and/or physical power. In other forms of negative social interaction, such as what adolescents often describe as “drama” (Marwick & boyd, 2011), intention and repetition may also be present, but the aggressive interactions are reciprocal and among equals. Thus, a key element distinguishing bullying from other aggressive behavior is a power differential between bully and victim (Olweus, 1978; Vaillancourt, Hymel, & McDougall, 2003). Vaillancourt and colleagues (2003) used a longitudinal survey of 6th to 10th graders to show the strong association between perceived power and bullying behaviors both among bullies and victims. Bullies target those with less power because targeting a weak victim elicits less social disapproval and also less self-imputation or self-examination, as the aggressive behavior might seem justified by the weakness of the victim (Bandura, 1978).

As Milner (2004) states, adolescents “have one crucial kind of power: the power to create an informal social world in which they evaluate one another” (p. 4). Evolutionary theories of aggression among adolescents argue that bullying is prevalent because it works to help bullies gain power and status (Ellis et al., 2012). While linked, power and status are distinct concepts. Power is “asymmetric control over valued

resources” (Magee & Smith, 2013, p. 159). Social power both leads to and results from social status, defined as popularity, liking, or social acceptance (Vaillancourt et al., 2003). Bullying can thus be considered both a display of and a reification of power with the social goal of a gain in status.

An asymmetry in power can be constituted in many ways, including physical strength, social status, or age. Weakness or difference among victims of bullying might be defined in terms of physical appearance, weight, emotional status, sexual promiscuity, or status as an outsider, based on social position, race or ethnicity, sexual orientation, or a combination of these characteristics (Englander, 2007). Aggression may also be a way to target and punish perceived violation from social norms, and those who inhabit identities with less social power may be more at risk for aggression or harassment (Duggan, 2017).

As with offline bullying, cyberbullying is also defined by a perceived power imbalance between aggressors and victims (Langos, 2012). However, the construct of power in online space is more difficult to define. While difference in physical power or size could be reflective of a power imbalance in a traditional bullying situation, online factors like technological expertise and anonymity can influence both aggressors and victims’ perception of power imbalance (Langos, 2012). Power differentials may be enabled or constructed by the structural features, or affordances, of social media. Valkenburg and Peter (2011) define the affordances of the Internet relevant to adolescents as anonymity, accessibility, and asynchronicity, all of which have consequences both for aggressors and victims. Asynchronicity means that bullies can post comments without seeing a victim’s reaction, a feature that is theorized to promote disinhibition. The asynchronicity of online communication also means that bullies can find their victims at any time via digital media and that victims can respond at any time. Permanence or accessibility means that interactions are visible to a potentially broad audience within a social network, and they can also be forwarded easily, at the click of a button.

Anonymity in particular is discussed as an affordance that may enable cyberbullying and could contribute to an imbalance of power between victims and aggressors. In one experiment, participants were more likely to post threatening messages if they knew their posts were anonymous (Lapidot-Lefler & Barak, 2012). Still, it is important to note that the anonymity in this case was bidirectional: participants were anonymous, but so were the other users they threatened. In instances of cyberbullying among adolescents, anonymity can be unidirectional. The aggressor knows the identity of the profile owner, but the aggressor’s identity is masked, both to the victim and to the rest of the networked public. The ability to attack anonymously might enable or even promote bullying through disinhibition and a perceived lack of negative consequences (Englander & Muldowney, 2007; Vandebosch & Van Cleemput, 2008). Adolescents associate not knowing their online attackers with feelings of powerlessness; not

knowing the aggressor's identity made it difficult or impossible for victims to place the attacks within a social context (Huitsing, Veenstra, Sainio & Salmivalli, 2012; Vandebosch & Van Cleemput, 2008).

Online victims of bullying cannot identify their aggressor 46% of the time, compared with 12% of the time offline (Ybarra et al., 2012). Still, victims can be reasonably sure that they know even an anonymous or unnamed attacker as many ties on social media are among those who know each other offline. The shield of anonymity online is a potential method of establishing a power differential where none exists in the offline social world or even in flattening an offline power difference. Adolescents suggest that anonymity can empower those who couldn't be bullies offline, due to lack of physical or social power, to be aggressors on social media (Vandebosch & Van Cleemput, 2008).

In cyberbullying as in relational aggression, attempts to demonstrate power and accrue status through aggression may be enacted through platform affordances, such as the number of friends on a social network site (Slonje, Smith, & Frisén, 2013), or through discourse rather than physical attacks. The focus of this study is on signaling power in aggressive discourse. Discourse approaches to social psychology posit that talk and text serve social functions (McKinlay & McVittie, 2008). While interviews or self-report surveys capture descriptions of behavior, discourse captures how language is used to negotiate social goals. For example, an ethnographic study of adolescent girls documents how discourse is used to maintain group boundaries (Goodwin, 2002b). Girls within a group identified as high-status or popular excluded others by depicting them as having no friends or of being of a lower social class (Goodwin, 2002a) or by claiming they are too sexualized or don't adhere to standards of beauty and thinness (Currie, Kelly, & Pomerantz, 2007). In addition, because power or lack of it is related to adherence to group norms, standards of what determines being slutty, overweight, or pretty are not based on concrete standards but on surveillance by peers (Currie et al., 2007). Thus, the power asymmetries fundamental to social exclusion and bullying are constructed in part via discourse.

We define discursive strategies as the language used by online aggressors, victims, and bystanders to accomplish social goals. Our first aim is to investigate discourse used by aggressors:

RQ1. What discursive strategies do online aggressors use most frequently to demonstrate social power over victims?

Cyberbullying and Victim Roles

In contrast to goals of aggressors in exerting social and physical power over victims, the goal of cyberbullying victims is protective. Protection of self and status can be achieved through several behaviors, both passive and provocative

(Olweus, 1978). Offline, victims respond to bullying by ignoring aggression, asking friends or adults for help, telling the bully to stop, fighting back, or running away. The affordances of the Internet create a different environment for interaction, however, in which victim strategies might also differ. Patchin and Hinduja (2006) surveyed victims of cyberbullying and found that 25% reported they did nothing while 15%–35% told bullies to stop. Smith et al. (2008) found that cyberbullying victims responded to bullying by doing nothing, reporting to offline individuals, reporting to the site administrators, asking the bully to stop, or fighting back against the cyberbully. While past research has cataloged victim responses through self-report, this research uses text in situ to categorize victim responses to cyberbullying. Because discourse is the focus of this study, we focus on behaviors that can be observed through text, such as asking the bully to stop or becoming aggressive against the bully. We also ask whether different strategies of attack predict different victim responses:

RQ2. Which strategies do cyberbullying victims use most frequently to respond to aggressive comments?

RQ3. How do cyberbullying victims' responses vary as a function of the type of aggressive behavior?

Cyberbullying and Bystanders

It is important to recognize that bullying is also embedded in a social context. Bullying has been conceptualized as not just a dyadic interaction but also a group interaction in which bystanders as well as victims and bullies have defined roles (Salmivalli, Lagerspetz, Bjorkqvist, & Österman, 1996). The participant roles approach outlined by Salmivalli and colleagues (1996) identifies roles of adolescents beyond the bully–victim dyad. The audience of face-to-face bystanders not actively involved in bullying may act in one of three ways: actively defend the victim, reinforce bullying activity by laughing or encouraging the aggressor, or reinforce passively by doing or saying nothing (Sutton & Smith, 1999).

Anti-bullying interventions emphasize the roles of bystanders in bullying encounters. Offline, bystanders can intervene actively by resisting the bully, although the ability to intervene is limited by time and space. Online, these considerations are relaxed, and the affordances of the Internet create a potentially limitless audience for cyberbullying. Bystanders can also respond anonymously or asynchronously. It may be the case that bystanders believe that the situation is resolved and there is no need for a response if an aggressive interaction is viewed after its initial posting (Allison & Bussey, 2016). When bystanders do decide to intervene online, it is often because they believe the victim will benefit from getting support from others after a cyberbullying incident (DeSmet et al., 2016).

Through in-depth interviews with adolescents, Marwick and boyd (2011) found that a perceived goal of adolescent victims is actually to elicit positive feedback from bystanders or to gain attention, a goal that begs the question of how common bystander intervention is online and what discursive strategies are used by bystanders. To this end, we ask the following questions:

RQ4. Which strategies do bystanders use to intervene in online aggression?

RQ5. How do bystander strategies vary as a function of (a) aggressors' attacks and (b) victims' responses?

In an analysis of aggressive interactions on the social media site Formspring, Moore, Nakano, Enomoto and Suda (2012) found that bystander defense of victims sometimes served to neutralize the power imbalance. Online, bystanders can intervene asynchronously and anonymously as well, which may encourage intervention or even alter the characteristics of interventions. However, it is not yet known how the tone of bystander responses, neutral, aggressive, or supportive, may promote or deter additional aggression. Finally, we ask the following question:

RQ6. Does the tone of bystander responses predict the likelihood of additional aggressive comments?

Method

Sample

The social media site ask.fm was selected for this study because of the high proportion of adolescent users (more than 50% under the age of 18 at the time data were collected) (Van Grove, 2013). In addition, the default privacy setting for profiles is to make content public, thus enhancing Information and Communication Technology (ICT) affordances that contribute to the emergence of cyberbullying. Also, the site's format requires that questions be answered before they are posted to a profile owner's page, so all aggressive comments would require at least a one-word response from the profile owner to be posted to the profile page. Thus, we could analyze patterns within aggressive interactions between a profile owner and an aggressor. In conversation analysis, a directive and paired response are called adjacency pairs (Schegloff & Sacks, 1973).

We used a two-stage process to collect data. First, we identified accounts containing aggressive content that could be classified as cyberbullying according to the literature. Accounts were identified using Google's site-specific search function. A snapshot of accounts that returned hits for the search terms was collected on 3 October 2013. This "snapshot" approach has been used in other content analyses of user-produced content that rely on search

engines to return results and ensures that the sample remains consistent, since web content is known to change over time (Taraszow, Aristodemou, Shitta, Laouris, & Arsoy, 2010).

The search terms used were a string of six phrases collected from news accounts of online and offline bullying and from searching the site. This set of terms included the following: kill yourself, die, waste of space, everyone hates you, you suck, and no one likes you. A profile was considered for the sample if it returned a post that contained at least one of the search phrases that was clearly aggressive. While we could not assess user intent, in determining what was aggressive, we considered not just the result initially returned but context and previous and subsequent posts. We collected profile information and links to aggressive posts returned from the Google search for 100 profiles that met the sample requirements. Of the initial profiles collected, 78 remained in the study after removing 10 deactivated accounts, 11 accounts that had removed the original aggressive post and all other aggressive questions, and 1 account suspended by the site.

Next, we collected posts from the 78 profiles identified in Stage 1 of the sampling protocol. We loaded all posts 6 months prior to the initial post retrieved from search results. This was done to provide an adequate timeframe from the point of the initial aggressive post returned in Stage 1, which varied by date for each user. The site does not provide exact dates for each post but instead categorizes the posts only by month and year. Additional aggressive posts were identified by searching the individual profile using the same set of keywords from the first stage of sampling or by researcher scan of the profile for additional aggressive posts that did not include the original keywords. We collected up to 11 aggressive post sequences from each of the 78 profiles.

The unit of analysis for this study was individual posts and responses. Once a post that met inclusion criteria was identified, it was copied into a database along with the response, poster information, one post after, response to one post after, two posts after, and response to two posts after. Posts and responses following the initial question were included in order to observe bystander reactions to cyberbullying. A total of 352 initial aggressive posts and their responses were collected, which yielded an additional 641 question and answer sets that occurred after the initial questions. In some instances, there were no posts after the original aggressive posts.

As noted earlier, posts to the site are public by default, although users do have the option to make profiles private. To ensure user privacy, we removed the URL indicating each profile from the data file after coding was completed. In addition, we have not quoted directly from any posts so that specific profiles cannot be identified via additional web searching (boyd, 2015). We did not attempt to contact profile owners in any way.

Pilot Study

We conducted a brief pilot study to refine coding categories for our variables of interest. Twenty profiles were randomly selected for qualitative analysis. Based on this in-depth analysis of aggressive social media interactions, we developed a typology of discursive strategies through which bullies demean victims within individual interactions, operationalized in this study as a question–answer dyad. We found that, in addition to anonymity, online aggressors attempt to demonstrate their victims' weakness through attacks on social status, appearance, sexual orientation, and personality and also by exhorting victims to commit suicide. In our pilot analysis of aggressive interactions, victims sought to neutralize the power dynamic by brushing off attacks via neutrality or sarcasm, attacking or demeaning the aggressor so that his or her claims seem unimportant, and countering the aggressors' claims.

We studied the comments posted after a social media attack to investigate whether the additional comments are neutral, aggressive, or supportive of the profile owner, the victim of the attack. We also investigated the specific strategies used by bystanders to either bolster the victim or denounce the attacker. These insights contributed to the development of coding categories as described below.

Operational Measures

Aggressive Posts. We coded elements of aggressive posts to assess discursive strategies used by aggressors to weaken victims. First, we coded whether aggressors posted anonymously or whether the statement included a link to the aggressor's profile. References to social status were defined as negative statements about the profile owner's social position or social exclusion, such as being disliked, hated, or excluded by a group (Vaillancourt et al., 2003). Examples include "no one likes you," "everyone hates you," or other references to popularity. An age variable was also coded as present or absent and was defined as references to the grade or age of a profile owner in a derogatory or negative way. A sexual orientation variable was counted as present if the poster included negative statements about the profile owner's sexual orientation or if they used anti-gay slurs. Moralizing statements regarding the sexual activities or over-sexualization of the profile owner, such as calling them a "slut" or "whore," were coded. Negative references to the attractiveness of a profile owner, for example, calling them ugly, were coded, as were statements attacking an individual for their weight, such as calling an individual fat or suggesting they need to lose weight. Exhortations to suicide were coded if the poster told the profile owner to kill or harm himself or herself or suggested any techniques for doing so, such as "cut yourself" or "drink bleach." Finally, references to someone being attention seeking or doing something just to receive attention from others were coded.

Profile Owner Responses. Profile owners' responses were coded for overall strategy and then specific sub-strategies used to respond to aggressive comments. A response was considered agreement if the content indicated that the owner agrees with the statements made in the aggressive posts, such as stating "I know" or "I will." Sarcasm was coded if the response seemed to mock the aggressive post or the poster, such as "I'll get right on that," after receiving a post telling the profile owner to kill themselves. Because sarcasm is contextual, coders assessed sarcasm in responses by considering the response in relation to the initial aggressive post. Neutral responses were coded if the poster replied with "ok" or other non-emotional comment. Counter-aggression was coded if the response attacks the aggressor with harsh language, profanity, or threats against them. Additional sub-strategies are described in Table 1.

Questions After Initial Post. Posts after the initial aggressive post were separated into three categories: aggressive, neutral, and supportive. A post was considered aggressive if it met the initial requirements outlined for sample collection in terms of containing one of the search keywords or clearly containing aggressive language as described in the bullying literature. Neutral posts were coded as present if the tone or language is informational, not emotionally charged, or questions unrelated to aggression. Supportive posts were defined as messages expressing support for the profile owner, either as a positive statement about them or as a counter-message aimed at the aggressive poster. If posts were coded as aggressive, the same variables previously described for the initial aggressive posts were coded. No further variables were coded if the post was considered neutral. Supportive posts were coded for several additional variables (Table 1).

Intercoder Reliability. Two coders were trained using the codebook and completed three rounds of intercoder training. Both coders independently coded 105 posts collected from 20 profiles randomly selected from the sample (Krippendorff's α coefficient, .71–1.0) (Krippendorff, 2004). The remaining posts were then divided and coded independently. The codebook listed variables and descriptions by post type: aggressive post, profile owner response, and supportive post. Descriptions of all variables and intercoder reliability statistics are presented in Table 1.

Results

Gender identity of profile owners was determined by a study of the profile page. Of the 78 profiles, 62 (79.5%) of profile pictures depicted females, 11 (14.1%) of profile pictures depicted males, and for 5 profiles (6.4%), gender could not be identified. The majority of profiles ($N=65$; 83.3%) were for individuals who appeared to be of high school age or

Table 1. Codebook and reliability.

Variable	Krippendorff's α coefficient	Variable definition	Example
Aggressive Posts			
Anonymous Poster	.85	No profile or name associated with aggressor	
Social status	.88	Negative statements about the profile owner's community status or social exclusion, such as being disliked or hated	"No one likes you," "You don't have any friends"
Age	1.00	References to the grade or age of a profile owner in a derogatory or negative way	"You're just a niner," "Everyone in grade 8 hates you"
Homosexuality	1.00	Negative statements about the profile owner being homosexual or anti-gay slurs	"Faggot," "Dyke," references to homosexual sex acts
Sexual Activities	.93	Moralizing statements regarding the sexual activities of the profile owner	"You're such a slut," "You look like a whore"
Attractiveness	.81	Negative references to the attractiveness of a profile owner	"You're ugly"
Weight	.79	Statements attacking an individual for their weight, such as calling an individual fat or suggesting they need to lose weight	"You're fat," "You should lose weight"
Attention Seeking	.84	Stating that the profile owner was behaving melodramatically to try to get attention.	"You're just an attention whore," "Stop crying just to get attention"
Suicide	.98	Exhortations to suicide were coded if the poster told the profile owner to kill himself or herself or suggested any techniques for doing so	"Kill yourself," "Drink bleach," "Die already"
Response			
Agreement	.87	Profile owner agrees with the statements made in the aggressive posts	"I know," "I will"
Sarcasm	.71	Clearly mocked aggressive post or the poster	"I'll get right on that," "You must be so proud of yourself"
Neutral	.84	If the poster replied with "ok" or other non-emotional comment	"Okay," "k"
Counter-aggression	.81	Response attacks the aggressor with harsh language, profanity, or threats against them	"Fuck off," "I'll fight you"
Quit	.87	Asking the poster to stop the aggression	"Leave me alone," "Stop"
Motive	.88	Questioning the poster's motive was included when profile owners asked why they would say that or other statements regarding why the aggression was taking place	"Why would you say that?"
Anonymous	.85	If the profile owner calls out the aggressor for being anonymous or "anon"	"Come off anon," "who are you?"
Hypothetical	.88	If the profile owner refers to situations that may occur, such as what the aggressive poster would do if the owner did commit suicide	"How would you feel if. . ."
Victim's Strength	.91	Statements mentioning the strength or resilience of the profile owner	"I can take it," "I'm strong"
Victim's Support	.92	Mentioning support from the profile owner's friends, family, or other social groups	"My friends love me," "I have lots of friends"
Victim's Identity	.90	Asserting identity as distinct from portrayal by questioner/aggressor	"That's not me," "You don't know me"
Minimization	.91	The profile owner used strategies to show that the aggression does not affect them	"I really don't care what you say"
Bullying	.92	References to bullying, such as calling the aggressor a bully or a hater	"You're a bully," "You're just a hater"
Additional posts			
Aggressive	.87	To be considered aggressive, posts had to meet the initial requirements outlined for sample collection	"You should kill yourself"
Neutral	.75	The tone or language is informational, not emotionally charged, or questions unrelated to aggression	"What's your favorite color?"

continued

Table 1. (Continued)

Variable	Krippendorff's α coefficient	Variable definition	Example
Supportive	.87	Messages expressing support for the profile owner, either as a positive statement about them or as a counter-message aimed at the aggressive poster	"We love you," "That anon is so stupid"
Supportive Posts			
Second Person	.78	The post was addressed directly to the profile owner	"You're so pretty," "You're awesome"
Ignore	.81	Ignoring statements, which include telling profile owners not to listen to the aggressor	"Just ignore that anon"
Care and Support	.78	Statements that affirm love or support for the profile owner	"I love you," "ily"
Third Person	.77	If the post describes the profile owner	"She is amazing," "She's great"
Counter-aggression	.83	The supporter attacked the aggressive poster with any threatening or questioning statements	"Why would you say that?" "You're a jerk!"
Anonymous	.80	Posts specifically calling out the anonymity of the aggressive poster	"Anons don't mean anything," "They wouldn't say that off anon"

younger or who claimed to be in high school in the profile description, whereas three profiles (3.8%) appeared to be for individuals older than high school age or who claimed to be older, and age could not be inferred for 10 profiles (12.8%).

Initial Aggressive Posts

RQ1 asked about the types of discursive strategies aggressors use most frequently against victims in this sample. Posts were a mean of 14.0 words long ($SD=15.5$, range = 1–63). Nearly all initial aggressive posts (97.2%) were posted anonymously, with only 2.8% of aggressors including their profile name in the aggressive post. The rhetorical strategies most commonly used in the sample to depict the target of aggressive posts as powerless were taunts directing the victim to commit suicide, attacks on the victim's social status, and pejorative terms related to the victim's sexual behavior, although coding for discursive strategies in initial aggressive posts was limited by the search terms used. In more than half of posts (54.3%), aggressors used language directing profile owners to kill themselves. Remarks about social status were found in 41.2% of posts. Aggressors often told profile owners that "No one likes you" or that the aggressor or another specific person "hated" or "couldn't stand" the profile owner. In more than a quarter of posts (26.7%), aggressors shamed profile owners with pejorative remarks related to sexual behavior (e.g., "whore" or "slut"). Only 6.0% of posts directly referred to conflicts regarding dating or sexual relationships. Figure 1 presents percentages for all aggressive tactics in initial aggressive posts as well as the additional aggressive posts coded in two subsequent questions.

Profile Owner Responses

RQ2 investigates the frequency of strategies used by profile owners/victims to respond to aggressive comments. There was substantial variation in the length of responses to initial aggressive posts. The mean length was 20.19 words ($SD=45.90$, range = 0–427). The strategy most commonly used to respond to aggressive posts was sarcasm, followed by counter-attacks, neutral responses, and then agreement. More than a third of profile owners (36.6%) responded to initial aggressive posts with sarcastic responses (e.g., to an aggressor posting "kill yourself," the profile owner says, "yeah, I'll get right on that"). Counter-aggression was nearly as common (31.3%) and ranged from brief expressions of profanity to more extensive attacks of several sentences. The next most frequent response type was neutral (19.6% of responses), which frequently entailed one-word responses such as "oh," "OK," or just ". . ." More infrequently, profile owners agreed with aggressors, without apparent sarcasm (8.0%). These responses included comments such as "I know no one likes me" or "I know I'm worthless."

In addition to these general response types, profile owners used other discursive strategies to respond. These strategies included asking aggressors directly to stop or to quit (9.7%), calling out aggressors for posting "hate" anonymously (9.7%), minimizing the effects of the aggression using statements such as "I don't care" or "it doesn't bother me" (8.0%), and refuting claims that the aggressor was depicting the profile owner's real identity (e.g., "you don't know me," 5.4%). A few participants also posed rhetorical questions interrogating how the aggressor would feel if the profile owner actually did commit suicide (3.7%). In only 3 of 352 responses did participants use the term "bully" (0.9%) (Figure 2).

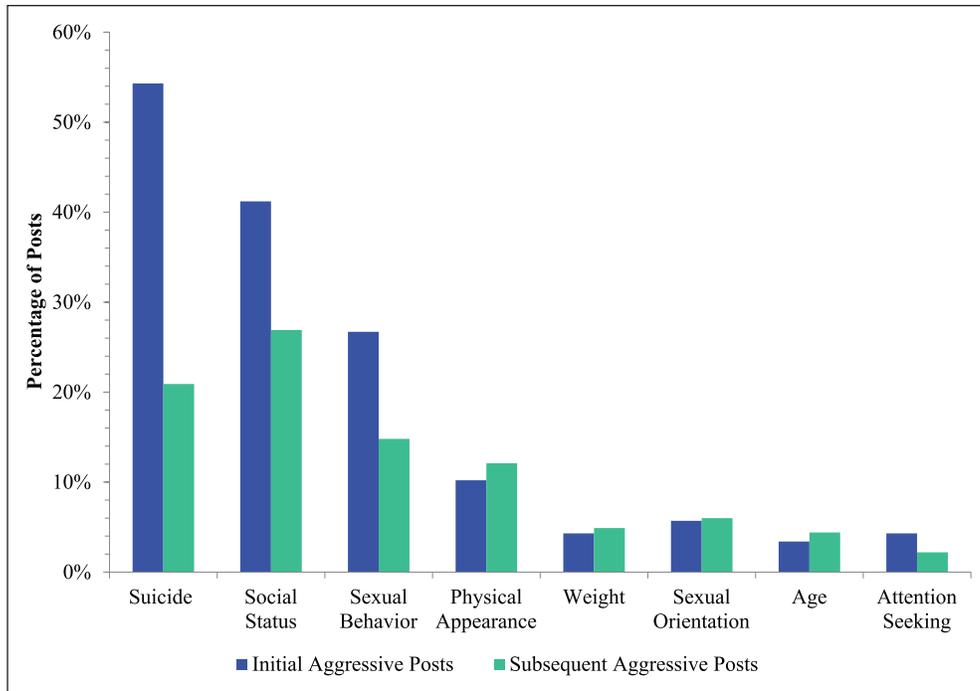


Figure 1. Use of discursive strategies in aggressive social media posts, % of posts.

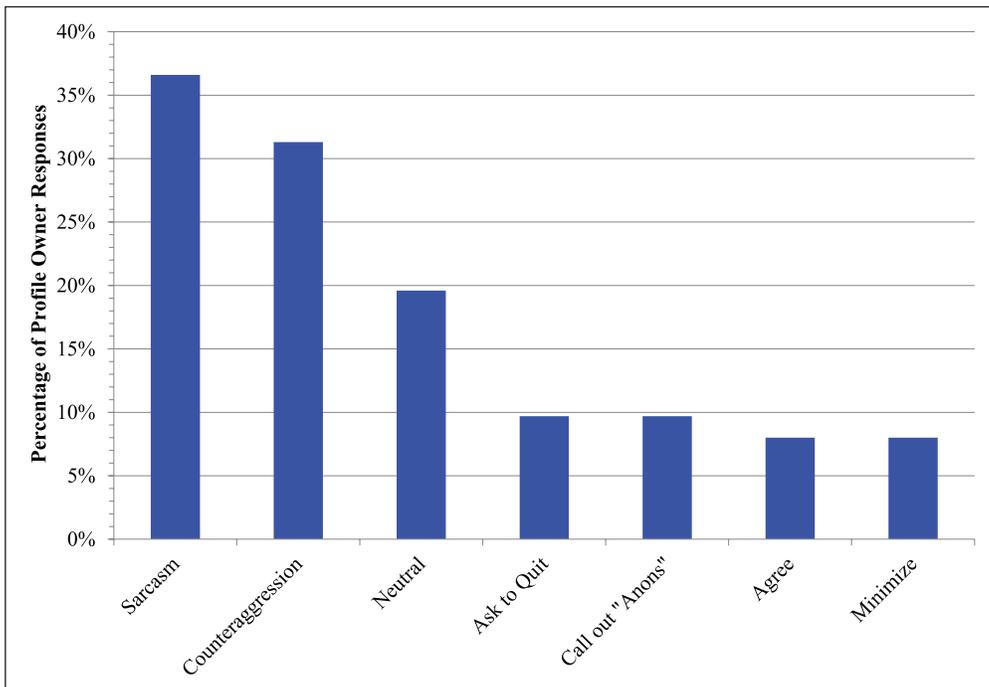


Figure 2. Most common discursive strategies in profile owner responses to aggressive posts, % of responses.

Next, we determined whether different attacks were associated with specific responses from victims (RQ3). When aggressors used pejorative terms regarding the profile owner’s sexuality, profile owners were significantly more likely to use counter-attacks, $\chi^2(1)=8.86, p=.003$. When pejorative

sexual terms were used in aggressive posts, profile owners were less likely to agree with aggressors in their responses, $\chi^2(1)=6.69, p=.01$. When aggressive posts implored profile owners to commit suicide, profile owners were much more likely to respond with neutral words or phrases (e.g., “OK,”

“Oh”), $\chi^2(1)=19.46, p<.001$. There were no other significant results found in χ^2 analyses comparing aggressive post strategies to response types.

Bystander Responses

In addition to initial aggressive posts and response, we coded the two post-and-response pairs that followed the initial aggressive posts. Because we coded two question–answer pairs following the initial post, the sample size for these responses was 641. The majority of responses were anonymous, 86.5% versus 13.4% for which posters provided the link to their own profile. We coded for three initial types among posts following initial aggressive posts: neutral, or no reference to aggressive posts; supportive of the profile owner, with direct reference to the attacks; and aggressive (RQ4). Almost half of responses (47.6%) were neutral, 28.0% were aggressive, and 25.9% were supportive of the profile owner.

Of the 78 profiles included in the analysis, 4 (5.1%) had no supportive posts. The mean number of supportive posts for all profiles, including those with none, was 8.35 ($SD=6.86$, range=0–22). Among the 144 supportive posts, we coded for several potential support strategies. Most bystanders (70.8%) used second person to address their supportive posts to the profile owner. A total of 17.4% of bystanders told profile owners not to pay attention to aggressors or to ignore aggressive posts, while 20.8% affirmed profile owners by saying they cared about them (e.g., “I love you!”). Fewer bystanders (19.4%) supported profile owners indirectly, by attacking the original aggressors. These posts were phrased in the third person, such as “Hey, anon, leave her alone.” Of all supportive posts, more than a third (34.0%) used profanity or other aggressive language to attack the aggressor, and 22.2% criticized the aggressor for posting attacks anonymously.

In the 182 aggressive responses that appeared one or two posts after the initial aggressive post, aggressors maligned profile owner social status in 26.9% of posts, suggested suicide in 20.9% of posts, used pejorative terms related to sexual activity in 14.8% of posts, attacked physical attractiveness in 12.1% of posts, used slurs related to sexual orientation in 6.0% of posts, mentioned participants being overweight in 4.9% of posts and young age in 4.4% of posts, and accused profile owners of seeking attention in 2.2% of posts.

RQ5 asked whether different phrases from aggressors or responses from victims were associated with different bystander responses. When aggressors posted that the profile owner should commit suicide, bystanders were less likely to respond with a neutral post that did not address the attack, $\chi^2(1)=4.31, p=.04$ (RQ5a). When aggressors posted exhortations to suicide, responses were more likely to also be aggressive, $\chi^2(1)=5.53, p=.02$. No specific language from victims predicted the type of bystander response.

In comparing responses of victims to initial aggressive posts and bystander responses, we found that when the

profile owner agreed with the aggressor, bystanders were less likely to respond with neutral posts, $\chi^2(1)=4.60, p=.03$ (RQ5b). Sarcastic responses from profile owners were more likely to be followed by neutral posts from bystanders, $\chi^2(1)=7.88, p=.005$. Counter-attacks toward aggressors from profile owners were more likely to be followed by supportive posts from bystanders, $\chi^2(1)=5.17, p=.02$. Profile owner references to repeated bullying, while infrequent, were also more likely to elicit supportive comments, $\chi^2(1)=11.09, p=.001$. Again, although profile owners responded infrequently to attacks by posing hypothetical scenarios (e.g., “How would you feel if I actually did kill myself?”), when profile owners did post those scenarios, they were less likely to be followed by additional aggressive posts, $\chi^2(1)=3.66, p=.05$. Finally, profile owner reference to repeated bullying (e.g., “I get this all the time”) was less likely to be followed by additional aggressive posts, $\chi^2(1)=6.37, p=.01$.

Next, we asked whether the tone of the first response by bystanders predicts the likelihood of an additional aggressive comment (RQ6). In general, the tone of the first bystander response predicted the second response. If the first bystander response was supportive, it was more likely to be followed by an additional supportive comment, $\chi^2(1)=18.77, p<.001$, and less likely to be followed by an additional aggressive comment, $\chi^2(1)=8.98, p=.003$. Conversely, if the first bystander response was aggressive, the second bystander response was much more likely to be coded as aggressive as well, $\chi^2(1)=47.47, p<.001$.

Discussion

The goal of the study is to investigate the discursive strategies used by aggressors, victims, and bystanders to negotiate issues of power and social status in aggressive interactions online. Overall, we found that language use varied by role within the encounter, although aggressors, victims, and bystanders all used aggressive language to attack and counter attack. In addition, we found some interesting linkages between discursive strategies that provide fertile ground for future studies and guidance for digital citizenship or antibullying interventions.

Anonymity is a key affordance of the site we examined, so it is not surprising that nearly all attackers remained anonymous. Previous research asserts that anonymity, and thereby the ability to deny being an attacker on social media, may encourage users to post aggressive messages online (Nilan, Burgess, Hobbs, Threadgold, & Alexander, 2015). What is more surprising is that supportive comments were often posted anonymously as well. While anonymity has been discussed as enabling aggression, future studies should examine in more detail whether anonymity also enables supportive comments from bystanders who might otherwise not have intervened in aggressive interactions. In addition, qualitative research with adolescents should investigate how anonymity

aligns with motivations to bully, since the veiled identity of attackers means that the online audience at least does not identify the bully's display of power with a particular individual. In other words, the power over the victim asserted in the aggressive comments does not clearly accrue to the attacker. Because demonstration of power is central to the definition of bullying and cyberbullying, it is important to understand what other factors motivate cyberbullying when the bully's identity is not known to bystanders.

In addition, while much of the research on anonymity and aggression in a computer-mediated communication (CMC) context often looks at the emergence of aggressive behavior as a function of ICT-afforded anonymity, less research looks at how anonymity of aggressors influences the victims and motivates bystander behavior. Online, there is often nuance to the concept of anonymity, especially when online interactions are taking place within members of an offline social network. Past researchers have shown that victims attacked online by anonymous aggressors report feeling more powerless because they cannot place these attacks within a social context (Vandebosch & Van Cleemput, 2008). However, even when an attacker's identity is not broadcast through a name or profile picture, victims could know or suspect the identity of an attacker based on past online or offline interactions. In this case, the twinning of public anonymity and the victim's private knowledge of aggressor identity or social group could lead to increased perception of severity by victims. Future research could pair interviews or surveys with screen captures of past aggressive encounters to add to our understanding of the range of ways in which anonymity may play out in aggressive interactions on social media.

One of our most compelling findings was that among victims and supportive bystanders, the term "anon" was used as a pejorative, to attack the attacker as too cowardly to reveal his or her identity. It is possible that being anonymous might have social costs as well as benefits. Calling out an attacker as "anon" was a way for profile owners to regain power after being attacked, by defining the attacker as too cowardly to say mean things face-to-face. Of interest, this discourse suggests that aggression face-to-face demonstrates power, whereas anonymous aggression demonstrates weakness. Interview research suggests that the function of anonymity in online adolescent social life is complex and that adolescents use anonymity strategically online to manage identity within sensitive encounters or circumvent social strata (Ellison, Blackwell, Lampe, & Trieu, 2016). Research on cyberbullying should continue to engage with the construct of anonymity to understand the nuanced role it plays in enabling aggression and responses to aggression as well as its influences on perceived severity of aggression on different platforms and within different social contexts. That said, it is also important not to focus on anonymity exclusively as the catalyst for aggression, since aggression is socially situated and informed by social structures that disempower or denigrate certain groups. As

Phillips (2015) argues in her work on trolling, anonymity may enable certain forms of aggression, but it is not a necessary and sufficient condition or a sole cause of aggressive behavior online.

In general, reciprocal online aggression is more common than cyberbullying, which is characterized by a power differential between the bully and victim (Law, Shapka, Domene, & Gagné, 2012). It was beyond the scope of this study to classify coded interactions as cyberbullying or reciprocal online aggression, but anonymity might serve different functions within bidirectional and unidirectional aggressive encounters. Future research should also investigate whether other affordances, namely, asynchronicity and accessibility, have any positive consequences for victims in addition to the negative consequences frequently mentioned in media reports. Asynchronicity could potentially provide more time to draft a response to aggressors, while accessibility could provide opportunities for bystanders to intervene.

In the online interactions included in this study, it is impossible to determine the identity of participants or the social context for aggressive encounters. Thus, while most of the profiles had more than one aggressive post, and sometimes many more, we could not confirm absolutely that these interactions met the repetition and power difference conditions for cyberbullying. As stated earlier, however, we could consider the language expressed within each type of post, and aggressive posts did demonstrate several significant ways in which online attackers sought to portray their victims as weak and thus to articulate a power differential. Aggressors very often attacked social status directly, telling victims that no one liked them, and they also frequently suggested that victims erase themselves entirely by committing suicide, suggesting that everyone would be better off. The vitriol of the language used to demean and degrade the victim was consistent across aggressive posts. These interactions are of great public health concern. Past research (Van Geel, Vedder, & Tanilon, 2014) showed positive associations between cyberbullying perpetration and victimization and suicide ideation and attempts. Repeated and highly aggressive attacks, like suicide-related calls-to-action, might pose significant health threats to individuals who are already vulnerable and at risk of suicide. In our sample, victims were likely to respond to exhortations to suicide with non-committal neutral comments, rather than the sarcasm or counter-attacks that could be seen as more active attempts to regain social power, suggesting perhaps that victims were unsure how to respond. Given the frequency of comments about suicide and self-harm, antibullying interventions could include more information about suicide or more concrete suggestions about how victims should respond. In our sample, hypotheticals that prompted aggressors to reflect on their attacks, while used rarely, were more effective at halting additional aggression

attempts. This preliminary finding can also be investigated in future research to determine whether this sort of hypothetical scenario is effective at forestalling online aggression or prompting bystander intervention.

In keeping with the gender imbalance in cyberbullying identified in our and other studies (e.g., Festl & Quandt, 2013), the next most frequent method used to demean or weaken victims was an attack on sexual behavior, either through pejorative terms or through direct reference to sex acts supposedly enacted by the victim. The prevalence of these attacks aligns with some studies of bullying among girls that suggest a primary goal of the aggression is to malign others perceived as direct competition for romantic partners (e.g., Ellis et al., 2012). Exposure to such aggressive sexual commentary has been found to correlate with poor body image and other mental health problems among adolescent boys and girls (Tiggemann & Slater, 2013). It is interesting to note that these attacks also prompted the most counter-attacks by victims, suggesting that attacks on sexuality may be felt particularly keenly among adolescent girls.

A strength of our study was a focus on actual victim responses rather than self-reported responses. Of great interest, sincere responses demonstrating hurt were rare. This aligns with suggestions that adolescent victims of aggression often seek to downplay attacks or avoid words like bullying to save face (Marwick & Boyd, 2011). Instead, victims often sought to actively counter aggressors either by shrugging off or minimizing the effects of attacks, through sarcasm or brief neutral responses, or through counter-attacks, in which aggressors were in turn demeaned as pathetic or cowardly, for failing to identify themselves. In the self-report literature on victim responses, categories are less fine-grained and may refer to talking back to the aggressor but not the specific strategies used, such as sarcasm, which may allow victims to regain social power, an opportunity that might not be afforded by ignoring or reporting. Providing specific language victims could use may increase the efficacy of victims or bystanders to respond in ways that shut down aggression without loss of face for victims or without fanning the flames of reciprocal aggression with subsequent posts.

One salient aspect of ask.fm in particular is that profile owners have the ability to block anonymous comments or simply to not answer aggressive questions. Since questions are only posted once they are answered, ignoring aggressive questions would be one way to avoid engaging in aggressive interactions. In a Pew survey, half of adolescents report deleting someone else's comment from their profile (Madden et al., 2013). While some aggressive comments from our initial search were deleted, many remained, months after the original search. In interviews, adolescents state that profile owners allow aggressive posts as a way of seeking attention or supportive comments (Marwick & Boyd, 2013), and "attention seeker" was a charge sometimes levied at profile owners. Similarly, social media

profiles that disclosed more personal information often evoked less empathy and more victim blaming when they received cyberbullying responses to their posts (Schacter, Greenberg, & Juvonen, 2016). This finding might relate to the "attention seeker" view of profiles owners who allow aggressive posts on their profiles even though they are able to block them. Future research should investigate in more detail motivations by victims to engage in aggressive interactions online. The fact that individuals seem to actively seek an audience for these interactions is interesting for discussions of identity management online.

One motivation for allowing aggressive comments to be seen publicly is to elicit support. While all but four profiles in our sample had at least one supportive post, unrelated and aggressive follow-up posts were more common. Still, when a supportive post followed the initial interaction, subsequent supportive posts, especially when phrased as comfort for the victim, were more likely. This suggests that intervening right away after an online attack could interrupt the aggression in such a way that future attacks are less likely. While most adolescents do not identify as victims of cyberbullying, many more have likely witnessed cyberbullying of friends or peers. Our findings are preliminary and not generalizable to other sites and situations, but if future research confirms that supportive posts from bystanders stall aggression, online civility training for adolescents should include advice about when and how to respond to support victims. Suggestions for bystander intervention based on evidence from online discourse can increase efficacy among bystanders to cyberbullying, who may feel that they should respond but aren't sure how.

Finally, since data were collected for this study, ask.fm has lost some cache among adolescent users, who likely migrated to other sites as the notoriety of ask.fm grew among parents and other adults. While adolescents may consider anonymity desirable in some situations (Ellison et al., 2016), social media sites where anonymity is required or even just the norm have come under attack. Recent media coverage warns parents that their children may experience cyberbullying when on sites that allow anonymous interactions, such as ask.fm (McGinn, 2016). Ask.fm addressed aggressive online interactions by forming a Safety Advisory Board to make the platform a "safer and more positive place" (Ask.fm, 2015, para 3). Despite these changes, the company has changed ownership twice since 2014 (British Broadcasting Corporation [BBC], 2016). The widespread criticism of anonymous social media sites as fertile ground for aggressive posts and incidents of cyberbullying may outrage adults, but changes to the sites, such as increased monitoring or disallowing anonymity, may drive away adolescent users. The lure of anonymity is strong enough that even as sites, like Formspring, collapse, others pop up to take their place. Future research should investigate how anonymity and aggressive interactions may affect a platform from a business perspective.

Study Implications

This study offers a number of conceptual and practical implications related to the conduct of cyberbullying, victims' responses, and bystander intervention on a social media platform where users mostly post anonymously. First, aggression bred aggression. However, aggression from victims was not as effective at stalling aggression in our sample as other strategies, such as posing hypotheticals. From a practical perspective, educators and community activists should consider how to equip adolescents to respond to online aggression, beyond just telling them to ignore it or report it. Second, the role of bystanders, a subject of much recent research, should be considered not just in terms of action or inaction but in the type of action taken. As they are able to also post anonymously, bystanders engaged in counter-arguing and responding to aggressive comment with aggressive comments themselves and also by pulling focus from bullies and supporting victims. The effectiveness of these approaches among perpetrators and victims of cyberbullying is yet to be fully investigated, yet they offer directions for future campaigns and interventions that leverage the power of bystanders, in terms of both responding to bullies and enhancing victims' resilience.

Limitations

This study is limited by the difficulties of sampling from social media. As stated, some profiles we initially identified were gone by the time we completed data collection. The snapshot nature of our sample also meant we cannot generalize from our analysis to all aggressive interactions or even aggressive interactions on this site. Future research should employ big data analytic procedures to gather more comprehensive samples and utilize machine learning techniques to detect instances of cyberbullying and aggressive online behaviors, as with the recent analysis of civility in comments from the *New York Times* (Muddiman & Stroud, 2017). This study is further limited by the nature of the social media site, which only posts the question-answer dyad once the profile owner answers the question. Naturally, we cannot claim to capture aggression that was not publicly responded to by the profile owner. In addition, we were not attempting to map out an aggressive interaction from start to finish online, so our initial posts were likely not the first aggressive post to have been fielded by each profile owner. The site has users across the globe, and we could not limit our study by geographic location, since most profile owners do not provide their location. A natural extension to this line of research is replicating our content analysis method with other platforms that vary in terms of affordances, especially in relation to anonymity. Finally, our search was limited by the initial terms used, selected based on a review of news articles about cyberbullying but by no means exhaustive, although we expanded the range of aggressive phrases by including subsequent aggressive posts. One major limitation of our initial search terms was a lack of homophobic, racist, or

misogynistic words or phrases. Because of this limitation, we may have missed instances of cyberbullying based on race/ethnicity, gender, or sexual orientation, all of which may be targets for online aggression. A more comprehensive list of keywords could aid future researchers in a more generalizable analysis of aggression discourse. Although online racism, sexual harassment, and cyberbullying have sometimes been treated as distinct constructs in the literature, understanding how race, gender, and sexual orientation put people at risk of online attacks or form the basis for aggressive rhetoric online is a crucial area for future scholarship.

It is also important to note that most vitriolic aggressive posts were followed by responses that seemed pedestrian by comparison. While we focus on aggressive posts in this analysis, such posts are by no means the most common on this or likely any other social media site. The analysis of Formspring found that 94% of all posts were neutral (Moore et al., 2012). Cyberbullying has gained prominence in the public agenda, but it bears repeating that much of what adolescents do on social media is likely not aggressive but a much less violent form of social negotiation or interaction.

Conclusion

This study is one of the first to investigate discursive strategies used in aggressive online interactions that would likely be classified as cyberbullying by a lay audience. While the study is exploratory, it suggests many areas for future research, particularly research related to the differing social motivations and social rewards involved with cyberbullying. Ellis et al. (2012) suggest that effective antibullying interventions must acknowledge that bullying is adaptive for aggressors. Intervening successfully requires altering the cost-benefit analysis, such that bullying becomes more socially costly than beneficial. Additional research into the effects of the discourse studied here can be used to inform interventions that suggest the most effective ways for victims to respond online or for bystanders to intervene to ultimately alter the social balance sheet for online aggression.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

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

ORCID iD

Rachel Young  <https://orcid.org/0000-0003-0228-2569>

References

- Allison, K. R., & Bussey, K. (2016). Cyber-bystanding in context: A review of the literature on witnesses' responses to cyberbullying. *Children and Youth Services Review*, *65*, 183–194.

- Ask.fm. (2015, January). *Ask.fm forms first ever Safety Advisory Board*. Retrieved from <http://about.ask.fm/ask-fm-forms-first-ever-safety-advisory-board/>
- Bandura, A. (1978). Social learning theory of aggression. *Journal of Communication*, 28, 12–29.
- boyd, D. (2015). Making sense of teen life: Strategies for capturing ethnographic data in a networked era. In E. Hargittai & C. Sandvig (Eds.), *Digital research confidential: The secrets of studying behavior online* (pp. 79–102). Cambridge: The MIT Press.
- British Broadcasting Corporation (BBC) (2016, July 4). Ask.fm changes hands once again. *BBC Technology News*. Retrieved from <http://www.bbc.com/news/technology-36702766>
- Currie, D. H., Kelly, D. M., & Pomerantz, S. (2007). “The power to squash people”: Understanding girls’ relational aggression. *British Journal of Sociology of Education*, 28, 23–37.
- DeSmet, A., Bastiaensens, S., Van Cleemput, K., Poels, K., Vandebosch, H., Cardon, G., & De Bourdeaudhuij, I. (2016). Deciding whether to look after them, to like it, or leave it: A multidimensional analysis of predictors of positive and negative bystander behavior in cyberbullying among adolescents. *Computers in Human Behavior*, 57, 398–415.
- Duggan, M. (2017). Online harassment 2017. Pew Research Center. Retrieved from <http://www.pewinternet.org/2017/07/11/online-harassment-2017/>
- Ellis, B. J., Del Giudice, M., Dishion, T. J., Figueredo, A. J., Gray, P., Griskevicius, V., . . . Wilson, D. S. (2012). The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology*, 48, 598–623.
- Ellison, N. B., Blackwell, L., Lampe, C., & Trieu, P. (2016). “The question exists but you don’t exist with it”: Strategic anonymity in the social lives of adolescents. *Social Media + Society*, 2. Advance online publication. doi:2056305116670673
- Englander, E. (2007). Is bullying a junior hate crime? Implications for interventions. *American Behavioral Scientist*, 51, 205–212.
- Englander, E., & Muldowney, A. M. (2007). Just turn the darn thing off: Understanding cyberbullying. In D. L. White, B. C. Glenn, & A. Wimes (Eds.), *Persistently safe schools: The 2007 National Conference on Safe Schools* (pp. 83–92). Washington, DC: Hamilton Fish Institute.
- Festl, R., & Quandt, T. (2013). Social relations and cyberbullying: The influence of individual and structural attributes on victimization and perpetration via the Internet. *Human Communication Research*, 39, 101–126.
- Goodwin, M. H. (2002a). Building power asymmetries in girls’ interaction. *Discourse & Society*, 13, 715–730.
- Goodwin, M. H. (2002b). Exclusion in girls’ peer groups: Ethnographic analysis of language practices on the playground. *Human Development*, 45, 392–415.
- Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research*, 14, 206–221.
- Huitsing, G., Veenstra, R., Sainio, M., & Salmivalli, C. (2012). “It must be me” or “It could be them?” the impact of the social network position of bullies and victims on victims’ adjustment. *Social Networks*, 34, 379–386.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: SAGE.
- Langos, C. (2012). Cyberbullying: The challenge to define. *Cyberpsychology, Behavior and Social Networking*, 15, 285–289.
- Lapidot-Lefler, N., & Barak, A. (2012). Efforts on anonymity, invisibility, and lack of eye contact on toxic online disinhibition. *Computers in Human Behavior*, 28, 434–443.
- Law, D. M., Shapka, J. D., Domene, J. F., & Gagné, M. H. (2012). Are cyberbullies really bullies? An investigation of reactive and proactive online aggression. *Computers and Human Behavior*, 28, 664–672.
- Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., Smith, A., & Beaton, M. (2013). Teens, social media, and privacy. *Pew Research Center Report*. Retrieved from <http://pewinternet.org/Reports/2013/Teens-social-media-and-privacy.aspx>
- Magee, J., & Smith, P. (2013). The social distance theory of power. *Personality and Social Psychology Review*, 17, 158–186.
- Marwick, A., & boyd, D. (2011, September 21–24). *The drama! Teen conflict, gossip, and bullying in networked publics*. Paper presented at A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, Oxford, UK.
- McGinn, D. (2016, January 21). Your kid on Ask.fm? Be afraid, very afraid. *The Globe and Mail*. Retrieved from <http://www.theglobeandmail.com/life/parenting/your-kid-is-on-askfm-be-afraid-very-afraid/article28308222/>
- McKinlay, A., & McVittie, C. (2008). *Social psychology and discourse*. Chichester, UK: Wiley-Blackwell.
- Milner, M., Jr. (2004). *Freaks, geeks, and cool kids: American teenagers, schools, and the culture of consumption*. New York, NY: Routledge.
- Moore, M. J., Nakano, T., Enomoto, A., & Suda, T. (2012). Anonymity and roles associated with aggressive posts in an online forum. *Computers in Human Behavior*, 28, 861–867.
- Muddiman, A., & Stroud, N. J. (2017). News values, cognitive biases, and partisan incivility in comment sections. *Journal of Communication*, 57, 586–609. doi:10.1111/jcom.12312
- Nilan, P., Burgess, H., Hobbs, M., Threadgold, S., & Alexander, W. (2015). Youth, social media, and cyberbullying among Australian youth: “Sick friends.” *Social Media + Society*, 1. Advance online publication. doi:10.1177/2056305115604848
- Olweus, D. (1978). *Aggression in the schools: Bullies and whipping boys*. Oxford, UK: Hemispheres.
- Patchin, J. W., & Hinduja, S. (2006). Bullies move beyond the schoolyard: A preliminary look at cyberbullying. *Youth Violence and Juvenile Justice*, 4, 148–169.
- Phillips, W. (2015). *This is why we can’t have nice things: Mapping the relationship between online trolling and mainstream culture*. Cambridge: MIT Press.
- Salmivalli, C., Lagerspetz, K., Bjorkqvist, K., & Österman, K. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, 22, 1–15.
- Schacter, H. L., Greenberg, S., & Juvonen, J. (2016). Who’s to blame? The effects of victim disclosure on bystander reactions to cyberbullying. *Computers in Human Behavior*, 57, 115–121.
- Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8, 289–327.
- Slonje, R., Smith, P. K., & Frisén, A. (2013). The nature of cyberbullying, and strategies for prevention. *Computers in Human Behavior*, 29, 26–32.
- Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry*, 49, 376–385.

- Smith-Spark, L. (2013, August 7). Hanna Smith suicide fuels calls for action on Ask.fm cyberbullying. *CNN.com*. Retrieved from <http://www.cnn.com/2013/08/07/world/europe/uk-social-media-bullying/>
- Sutton, J., & Smith, P. K. (1999). Bullying as a group process: An adaptation of the participant role approach. *Aggressive Behavior, 25*, 97–111.
- Tarasow, T., Aristodemou, E., Shitta, G., Laouris, Y., & Arsoy, Y. (2010). Disclosure of personal and contact information by young people in social networking sites: An analysis using Facebook profiles as an example. *International Journal of Media and Cultural Politics, 6*, 81–101.
- Tiggemann, M., & Slater, A. (2013). NetGirls: The Internet, Facebook, and body image concern in adolescent girls. *International Journal of Eating Disorder, 46*, 630–633.
- Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior, 26*, 277–287.
- Vaillancourt, T., Hymel, S., & McDougall, P. (2003). Bullying is power: Implications for school-based intervention strategies. *Journal of Applied School Psychology, 19*, 157–176.
- Valkenburg, P., & Peter, J. (2011). Online communication among adolescents: An integrated model of its attraction, opportunities, and risks. *Journal of Adolescence Health, 48*, 121–127.
- Vandebosch, H., & Van Cleemput, K. (2008). Defining cyberbullying: A qualitative research into the perceptions of youngsters. *Cyberpsychology & Behavior, 11*, 499–503.
- Van Geel, M., Vedder, P., & Tanilon, J. (2014). Relationship between peer victimization, cyberbullying, and suicide in children and adolescents: A meta-analysis. *JAMA Pediatrics, 168*, 435–442.
- Van Grove, J. (2013, June 8). Ask.fm, the troubling secret playground of tweens and teens. *CNet*. Retrieved from <http://www.cnet.com/news/ask-fm-the-troubling-secret-playground-of-tweens-and-teens/>
- Veenstra, R., Lindenberg, S., Zijlstra, B., De Winter, A. F., Verhulst, F. C., & Ormel, J. (2007). The dyadic nature of bullying and victimization: Testing a dual-perspective theory. *Child Development, 78*, 1843–1854.
- Ybarra, M., Boyd, D., Korchmaros, J. D., & Oppenheim, J. K. (2012). Defining and measuring cyberbullying within the larger context of bullying victimization. *Journal of Adolescent Health, 51*, 53–58.

Author Biographies

Rachel Young is an assistant professor of Journalism and Mass Communication at the University of Iowa. Her research investigates the role of social media and other user-generated digital content in public health, with a particular focus on online safety and cyberbullying.

Stephanie Miles is currently a communications associate at the Foundation for Food and Agriculture Research. Her research interests include digital media and health communication.

Saleem Alhabash is an assistant professor of Public Relations and Social Media at Michigan State University's Department of Advertising + Public Relations. His research focuses on the persuasive effects of digital and social media. More recently, his research has tackled issues related to alcohol marketing and advertising and digital aggression (cyberbullying).