

How to Stop the Bullies

THE ANGST AND IRE OF TEENAGERS IS FINDING NEW, SOMETIMES DANGEROUS EXPRESSION ONLINE—PRECIPITATING THREATS, FIGHTS, AND A SCOURGE OF HARRASSMENT THAT PARENTS AND SCHOOLS FEEL POWERLESS TO STOP. THE INSIDE STORY OF HOW EXPERTS AT FACEBOOK, COMPUTER SCIENTISTS AT MIT, AND EVEN MEMBERS OF THE HACKER COLLECTIVE ANONYMOUS ARE HUNTING FOR SOLUTIONS TO AN INCREASINGLY TRICKY PROBLEM.

By Emily Bazelon

In the annals of middle-school mischief, the Facebook page Let's Start Drama deserves an entry. The creator of the page—no one knew her name, but everyone was sure she was a girl—had a diabolical knack for sowing conflict among students at Woodrow Wilson Middle School in Middletown, Connecticut. “Drama Queen,” as I came to think of her in the months I spent reporting at the school to write [a book about bullying](#), knew exactly how to use the Internet to rile her audience. She hovered over them in cyberspace like a bad fairy, with the power to needle kids into ending friendships and starting feuds and fistfights.

In contrast with some other social networks, like Twitter, Facebook requires its users to sign up with their real names. Drama Queen easily got around this rule, however, by setting up Let's Start Drama with a specially created e-mail address that didn't reveal her identity. Wrapped in her cloak of anonymity, she was free to pass along cruel gossip without personal consequences. She started by posting a few idle rumors, and when that gained her followers, she asked them to send her private messages relaying more gossip, promising not to disclose the source. Which girl had just lost her virginity? Which boy had asked a girl to sext him a nude photo? As Drama Queen posted the tantalizing tidbits she gathered, more kids signed up to follow her exploits—a real-life version of *Gossip Girl*. She soon had an audience of 500, many drawn from Woodrow Wilson's 750 students, plus a smattering from the local high school and a nearby Catholic school.

Students didn't just message rumors to Drama Queen; they also commented in droves on her posts, from their own real Facebook accounts, or from other fake ones. As one kid wrote about Drama Queen on the Let's Start Drama page, “She just starts mad shit and most of the time so do the ppl who comment.”

Drama Queen was particularly ingenious at pitting kids against each other in contests of her own creation. She regularly posted photographs of two girls side by side, with the caption “WHOS PRETTIERRR?!” Below the pictures, commenters would heckle and vote. One such contest drew 109 comments over three days. When it became clear which contestant was losing, that girl wrote that she didn’t care: “nt even tryinqq to b funny or smart.” The rival who beat her answered, “juss mad you loss ok ppl voted me ! If you really loooked better they wouldve said you but THEY DIDNT sooo sucks for you.” This exchange nearly led to blows outside of school, other students told me. And they said a fight *did* break out between two boys who were featured on Let’s Start Drama, in dueling photos, above the caption “Who would win in a fight?” They reportedly ended up pummeling each other off school grounds one day after classes.

Melissa Robinson, who was a social worker for the [Middletown Youth Services Bureau](#), quickly got wind of Let’s Start Drama because, she says, “it was causing tons of conflict.” Robinson worked out of an office at Woodrow Wilson with Justin Carbonella, the bureau’s director, trying to fill gaps in city services to help students stay out of trouble. Their connecting suite of small rooms served as a kind of oasis at the school: the two adults didn’t work for the principal, so they could arbitrate conflict without the threat of official discipline. I often saw kids stop by just to talk, and they had a lot to say about the aggression on Let’s Start Drama and the way it was spilling over into real life. “We’d go on Facebook to look at the page, and it was pretty egregious,” Carbonella told me. Surfing around on Facebook, they found more anonymous voting pages, with names like Middletown Hos, Middletown Trash Talk, and Middletown Too Real. Let’s Start Drama had the largest audience, but it had spawned about two dozen imitators.

Carbonella figured that all of these pages had to be breaking Facebook's rules, and he was right. The site has built its brand by holding users to a relatively high standard of decency. "You will not bully, intimidate, or harass any user," Facebook requires people to pledge when they sign up. Users also agree not to fake their identities or to post content that is hateful or pornographic, or that contains nudity or graphic violence. In other words, Facebook does not style itself as the public square, where people can say anything they want, short of libel or slander. It's much more like a mall, where private security guards can throw you out.

Carbonella followed Facebook's procedure for filing a report, clicking through the screens that allow you to complain to the site about content that you think violates a rule. He clicked the bubbles to report bullying and fake identity. And then he waited. And waited. "It felt like putting a note in a bottle and throwing it into the ocean," Carbonella said. "There was no way to know if anyone was out there on the other end. For me, this wasn't a situation where I knew which student was involved and could easily give it to a school guidance counselor. It was completely anonymous, so we really needed Facebook to intervene." But, to Carbonella's frustration, Let's Start Drama stayed up. He filed another report. Like the first one, it seemed to sink to the bottom of the ocean.

Facebook, of course, is the giant among social networks, with more than 1 billion users worldwide. In 2011, *Consumer Reports* published the results of a survey showing that 20 million users were American kids under the age of 18; in an update the next year, it estimated that 5.6 million were under 13, the eligible age for an account. As a 2011 report from the Pew Internet and American Life Project put it, "Facebook dominates teen social media usage." Ninety-three percent of kids who use social-networking sites have a Facebook account. (Teens and preteens are also signing up in increasing numbers for Twitter—Pew found that 16 percent of 12-to-17-year-olds say they use the site, double the rate from two years earlier.)

Social networking has plenty of upside for kids: it allows them to pursue quirky interests and connect with people they'd have no way of finding otherwise. An online community can be a lifeline if, say, you're a gender-bending 15-year-old in rural Idaho or, for that matter, rural New York. But as Let's Start Drama illustrates, there's lots of ugliness, too. The 2011 Pew report found that 15 percent of social-media users between the ages of 12 and 17 said they'd been harassed online in the previous year. In 2012, *Consumer Reports* estimated that 800,000 minors on Facebook had been bullied or harassed in the previous year. (Facebook questions the methodology of the magazine's survey; however, the company declined to provide specifics.) In the early days of the Internet, the primary danger to kids seemed to be from predatory adults. But it turns out that the perils adults pose, although they can be devastating, are rare. The far more common problem kids face when they go

online comes from other kids: the hum of low-grade hostility, punctuated by truly damaging explosions, that is called cyberbullying.

What can be done about this online cruelty and combat? As parents try, and sometimes fail, to keep track of their kids online, and turn to schools for help, youth advocates like Robinson and Carbonella have begun asking how much responsibility falls on social-networking sites to enforce their own rules against bullying and harassment. What *does* happen when you file a report with Facebook? And rather than asking the site to delete cruel posts or pages one by one, is there a better strategy, one that stops cyberbullying before it starts? Those questions led me to the Silicon Valley headquarters of Facebook, then to a lab at MIT, and finally (and improbably, I know) to the hacker group Anonymous.

The people at Facebook who decide how to wield the site's power when users complain about content belong to its User Operations teams. The summer after my trips to Woodrow Wilson, I traveled to the company's headquarters and found Dave Willner, the 27-year-old manager of content policy, waiting for me among a cluster of couches, ready to show me the Hate and Harassment Team in action. Its members, who favor sneakers and baseball caps, scroll through the never-ending stream of reports about bullying, harassment, and hate speech. (Other groups that handle reports include the Safety Team, which patrols for suicidal content, child exploitation, and underage users; and the Authenticity Team, which looks into complaints of fake accounts.) Willner was wearing flip-flops, and I liked his blunt, clipped way of speaking. "Bullying is hard," he told me. "It's slippery to define, and it's even harder when it's writing instead of speech. Tone of voice disappears." He gave me an example from a recent report complaining about a status update that said "He got her pregnant." Who was it about? What had the poster intended to communicate? Looking at the words on the screen, Willner had no way to tell.

In an attempt to impose order on a frustratingly subjective universe, User Operations has developed one rule of thumb: if you complain to Facebook that you are being harassed or bullied, the site takes your word for it. "If the content is about you, and you're not famous, we don't try to decide whether it's actually mean," Willner said. "We just take it down."

All other complaints, however, are treated as "third-party reports" that the teams have to do their best to referee. These include reports from parents saying their children are being bullied, or from advocates like Justin Carbonella.

To demonstrate how the harassment team members do their jobs, Willner introduced me to an affable young guy named Nick Sullivan, who had on his desk a sword-carrying Grim Reaper figurine. Sullivan opened the program that he uses for sorting and resolving reports, which is known as the Common Review Tool (a precursor to the tool had a better name: the Wall of Shame).

Sullivan cycled through the complaints with striking speed, deciding with very little deliberation which posts and pictures came down, which stayed up, and what other action, if any, to take. I asked him whether he would ever spend, say, 10 minutes on a particularly vexing report, and Willner raised his eyebrows. “We optimize for half a second,” he said. “Your average decision time is a second or two, so 30 seconds would be a really long time.” (A Facebook spokesperson said later that the User Operations teams use a process optimized for accuracy, not speed.) That reminded me of Let’s Start Drama. Six months after Carbonella sent his reports, the page was still up. I asked why. It hadn’t been set up with the user’s real name, so wasn’t it clearly in violation of Facebook’s rules?

After a quick search by Sullivan, the blurry photos I’d seen many times at the top of the Let’s Start Drama page appeared on the screen. Sullivan scrolled through some recent “Who’s hotter?” comparisons and clicked on the behind-the-scenes history of the page, which the Common Review Tool allowed him to call up. A window opened on the right side of the screen, showing that multiple reports had been made. Sullivan checked to see whether the reports had failed to indicate that Let’s Start Drama was administered by a fake user profile. But that wasn’t the problem: the bubbles had been clicked correctly. Yet next to this history was a note indicating that future reports about the content would be ignored.

We sat and stared at the screen.

Willner broke the silence. “Someone made a mistake,” he said. “This profile should have been disabled.” He leaned in and peered at the screen. “Actually, two different reps made the same mistake, two different times.”

There was another long pause. Sullivan clicked on Let’s Start Drama to delete it.

With millions of reports a week, most processed in seconds—and with 2.5 billion pieces of content posted daily—no wonder complaints like Carbonella’s fall through the cracks. A Facebook spokesperson said that the site has been working on solutions to handle the volume of reports, while hiring “thousands of people” (though the company wouldn’t discuss the specific roles of these employees) and building tools to address misbehavior in other ways.

One idea is to improve the reporting process for users who spot content they don’t like. During my visit, I met with the engineer Arturo Bejar, who’d designed new flows, or sets of responses users get as they file a report. The idea behind this “social reporting” tool was to lay out a path for users to find help in the real world, encouraging them to reach out to people they know and trust—people who might understand the context of a negative post. “Our goal should be to help people solve the underlying problem in the offline world,” Bejar said. “Sure, we can take content down and warn the bully, but probably the most important

thing is for the target to get the support they need.”

After my visit, Bejar started working with social scientists at Berkeley and Yale to further refine these response flows, giving kids new ways to assess and communicate their emotions. The researchers, who include Marc Brackett and Robin Stern of Yale, talked to focus groups of 13- and 14-year-olds and created scripted responses that first push kids to identify the type and intensity of the emotion they’re feeling, and then offer follow-up remedies depending on their answers. In January, during a presentation on the latest version of this tool, Stern explained that some of those follow-ups simply encourage reaching out to the person posting the objectionable material—who typically takes down the posts or photos if asked.

Dave Willner told me that Facebook did not yet, however, have an algorithm that could determine at the outset whether a post was meant to harass and disturb—and could perhaps head it off. This is hard. As Willner pointed out, context is everything when it comes to bullying, and context is maddeningly tricky and subjective.

One man looking to create such a tool—one that catches troublesome material before it gets posted—is Henry Lieberman, a computer scientist whose background is in artificial intelligence. In November, I took a trip to Boston to meet him at his office in MIT’s Media Lab. Lieberman looked like an older version of the Facebook employees: he was wearing sneakers and a baseball cap over longish gray curls. A couple years ago, a rash of news stories about bullying made him think back to his own misery in middle school, when he was a “fat kid with the nickname Hank the Tank.” (This is hard to imagine now, given Lieberman’s lean frame, but I took his word for it.) As a computer guy, he wondered whether cyberbullying would wreck social networking for teenagers in the way spam once threatened to kill e-mail—through sheer overwhelming volume. He looked at the frustrating, sometimes fruitless process for logging complaints, and he could see why even tech-savvy adults like Carbonella would feel at a loss. He was also not impressed by the generic advice often doled out to young victims of cyberbullying. “‘Tell an adult. Don’t let it get you down’—it’s all too abstract and detached,” he told me. “How could you intervene in a way that’s more personal and specific, but on a large scale?”

To answer that question, Lieberman and his graduate students started analyzing thousands of YouTube comments on videos dealing with controversial topics, and about 1 million posts provided by the social-networking site Formspring that users or moderators had flagged for bullying. The MIT team’s first insight was that bullies aren’t particularly creative. Scrolling through the trove of insults, Lieberman and his students found that almost all of them fell under one (or more) of six categories: they were about appearance, intelligence, race, ethnicity, sexuality, or social acceptance and rejection. “People say there

are an infinite number of ways to bully, but really, 95 percent of the posts were about those six topics,” Lieberman told me.

Focusing accordingly, he and his graduate students built a “commonsense knowledge base” called **BullySpace**—essentially a repository of words and phrases that could be paired with an algorithm to comb through text and spot bullying situations. Yes, BullySpace can be used to recognize words like *fat* and *slut* (and all their text-speak misspellings), but also to determine when the use of common words varies from the norm in a way that suggests they’re meant to wound.

Lieberman gave me an example of the potential ambiguity BullySpace could pick up on: “You ate six hamburgers!” On its own, *hamburger* doesn’t flash cyberbullying—the word is neutral. “But the relationship between *hamburger* and *six* isn’t neutral,” Lieberman argued. BullySpace can parse that relationship. To an overweight kid, the message “You ate six hamburgers!” could easily be cruel. In other situations, it could be said with an admiring tone. BullySpace might be able to tell the difference based on context (perhaps by evaluating personal information that social-media users share) and could flag the comment for a human to look at.

BullySpace also relies on stereotypes. For example, to code for anti-gay taunts, Lieberman included in his knowledge base the fact that “Put on a wig and lipstick and be who you really are” is more likely to be an insult if directed at a boy. BullySpace understands that lipstick is more often used by girls; it also recognizes more than 200 other assertions based on stereotypes about gender and sexuality. Lieberman isn’t endorsing the stereotypes, of

course: he's harnessing them to make BullySpace smarter. Running data sets from the YouTube and Formspring posts through his algorithm, he found that BullySpace caught most of the insults flagged by human testers—about 80 percent. It missed the most indirect taunting, but from Lieberman's point of view, that's okay. At the moment, there's nothing effective in place on the major social networks that screens for bullying before it occurs; a program that flags four out of five abusive posts would be a major advance.

Lieberman is most interested in catching the egregious instances of bullying and conflict that go destructively viral. So another of the tools he has created is a kind of air-traffic-control program for social-networking sites, with a dashboard that could show administrators where in the network an episode of bullying is turning into a pileup, with many users adding to a stream of comments—à la Let's Start Drama. "Sites like Facebook and Formspring aren't interested in every little incident, but they do care about the pileups," Lieberman told me. "For example, the week before prom, every year, you can see a spike in bullying against LGBT kids. With our tool, you can analyze how that spreads—you can make an epidemiological map. And then the social-network site can target its limited resources. They can also trace the outbreak back to its source." Lieberman's dashboard could similarly track the escalation of an assault on one kid to the mounting threat of a gang war. That kind of data could be highly useful to schools and community groups as well as the sites themselves. (Lieberman is leery of seeing his program used in such a way that it would release the kids' names beyond the social networks to real-world authorities, though plenty of teenagers have social-media profiles that are public or semipublic—meaning their behavior is as well.)

I know some principals and guidance counselors who would pay for this kind of information. The question is what to do with it. Lieberman doesn't believe in being heavy-handed. "With spam, okay, you write the program to just automatically delete it," he said. "But with bullying, we're talking about free speech. We don't want to censor kids, or ban them from a site."

More effective, Lieberman thinks, are what he calls "ladders of reflection" (a term he borrowed from the philosopher Donald Schön). Think about the kid who posted "Because he's a fag! ROTFL [rolling on the floor laughing]!!!" What if, when he pushed the button to submit, a box popped up saying "Waiting 60 seconds to post," next to another box that read "I don't want to post" and offered a big X to click on? Or what if the message read "That sounds harsh! Are you sure you want to send that?" Or what if it simply reminded the poster that his comment was about to go to thousands of people?

Although Lieberman has had exploratory conversations about his idea with a few sites, none has yet deployed it. He has a separate project going with MTV, related to its Web and

phone app called [Over the Line?](#), which hosts user-submitted stories about questionable behavior, like sexting, and responses to those stories. Lieberman's lab designed an algorithm that sorts the stories and then helps posters find others like them. The idea is that the kids posting will take comfort in having company, and in reading responses to other people's similar struggles.

Lieberman would like to test how his algorithm could connect kids caught up in cyberbullying with guidance targeted to their particular situation. Instead of generic "tell an adult" advice, he'd like the victims of online pummeling to see alerts from social-networking sites designed like the keyword-specific ads Google sells on Gmail—except they would say things like "Wow! That sounds nasty! Click here for help." Clicking would take the victims to a page that's tailored to the problem they're having—the more specific, the better. For example, a girl who is being taunted for posting a suggestive photo (or for refusing to) could read a synthesis of the research on sexual harassment, so she could better understand what it is, and learn about strategies for stopping it. Or a site could direct a kid who is being harassed about his sexuality to resources for starting a Gay-Straight Alliance at his school, since research suggests those groups act as a buffer against bullying and intimidation based on gender and sexuality. With the right support, a site could even use Lieberman's program to offer kids the option of an IM chat with an adult. (Facebook already provides this kind of specific response when a suicidal post is reported. In those instances, the site sends an e-mail to the poster offering the chance to call the National Suicide Prevention Lifeline or chat online with one of its experts.)

Lieberman would like to build this content and then determine its effectiveness by asking kids for their feedback. He isn't selling his algorithms or his services. As a university professor, he applies for grants, and then hopes companies like MTV will become sponsors. He's trying to work with companies rather than criticize them. "I don't think they're trying to reflexively avoid responsibility," he told me. "They are conscious of the scale. Anything that involves individual action on their part, multiplied by the number of complaints they get, just isn't feasible for them. And it is a challenging problem. That's where technology could help a little bit. My position is that technology can't solve bullying. This is a people problem. But technology can make a difference, either for the negative or the positive. And we're behind in paying attention to how to make the social-network universe a better place, from a technological standpoint."

Internal findings at Facebook suggest that Lieberman's light touch could indeed do some good. During my visit to Silicon Valley, I learned that the site had moved from wholesale banishment of rule-breakers toward a calibrated combination of warnings and "temporary crippling of the user experience," as one employee put it. After all, if you're banished, you can sign up again with a newly created e-mail address under an assumed name. And you

might just get angry rather than absorb the message of deterrence. Instead, Facebook is experimenting with threats and temporary punishments. For example, the Hate and Harassment Team can punish a user for setting up a group to encourage bullying, by barring that person from setting up any other group pages for a month or two. (If the account associated with the offensive group uses a made-up name, then the site's only leverage is to remove the group.) According to an in-house study, 94 percent of users whose content prompted a report had never been reported to the site before. As Dave Willner, the content-policy manager, put it when he told me about the study: "The rate of recidivism is very low."

He explained, in his appealingly blunt way, "What we have over you is that your Facebook profile is of value to you. It's a hostage situation." This didn't surprise me. In the course of my reporting, I'd been asking middle-school and high-school students whether they'd rather be suspended from school or from Facebook, and most of them picked school.

The hacker group Anonymous isn't the first place most parents would want their bullied kids to turn. Launched a decade ago, Anonymous is best known for its vigilante opposition to Internet censorship. The group has defaced or shut down the Web sites of the Syrian Ministry of Defense, the Vatican, the FBI, and the CIA. Its slogan, to the extent a loosely affiliated bunch of hackers with no official leadership can be said to have one, is "When your government shuts down the Internet, shut down your government." Anonymous has also wreaked financial havoc by attacking MasterCard, Visa, and PayPal after they froze payments to the accounts of WikiLeaks, the site started by Julian Assange to publish government secrets.

Since Anonymous is anarchic, the people who answer its call (and use its trademark Guy Fawkes mask in their online photos) speak for themselves rather than represent the group, and protest in all kinds of ways. Some, reportedly, have not been kind to kids. There was the case, for example, of a 15-year-old named McKay Hatch, who started a No Cussing Club in South Pasadena, California. When the concept took off in other cities, a group referring to itself as Anonymous launched a countercampaign, No Cussing Sucks, and posted Hatch's name, photo, and contact information across the Web; he got 22,000 e-mails over two weeks.

But other people in Anonymous have a Robin Hood bent, and this fall, they rode to the rescue of a 12-year-old girl who'd come in for a torrent of hate on Twitter. Her error was to follow the feed of a 17-year-old boy she didn't know and then stop following him when he posted remarks she found rude. The boy took offense and, with three friends, went after her. The boys threatened to "gang bang" her, and one even told her to kill herself. "I'm gonna take today's anger and channel it into talking shit to this 12 year old girl," one wrote.

“Blow up [her Twitter handle] till she deletes her twitter,” another one added. The girl lived far from the boys, so she wasn’t in physical danger, but she was disturbed enough to seek help online. “I have been told to kill myself alot its scary to think people in the world want you to die :(,” she wrote to another Twitter user who asked me to call her Katherine. “He has deleted some of them he was saying things like do you have a rope? and didnt the bleach work?”

Her pleas reached Katherine in the wake of the suicide of a 15-year-old Canadian girl named Amanda Todd. Before Amanda died, she posted a video of herself on YouTube, in which she silently told her story using note cards she’d written on. Amanda said that a man she’d met online had persuaded her to send him a topless photo, then stalked her and released the photo, causing her misery at school. The video is raw and disturbing, and it moved Katherine and a member of Anonymous with the screen name Ash. “It made me choke up,” Ash told me. When Katherine discovered that people were still sending the compromising photo of Amanda around online, she and Ash teamed up to help organize a drive to stop them and report offending users to Twitter, which removes pornographic content appearing on its site.

As Katherine and Ash came across other examples of bullying, like rape jokes and suicide taunts, they found that “Twitter will suspend accounts even if they are not in violation of Twitter rules when simply 1000s of people mass report an account as spam,” Katherine explained to me in an e-mail. A Twitter spokesperson said this was possible (though he added that if spam reports turn out to be false, most accounts soon go back online). Twitter bans direct and specific threats, and it can block IP addresses to prevent users whose accounts are deleted from easily starting new ones. But the site doesn’t have an explicit rule against harassment and intimidation like Facebook does.

While monitoring Twitter for other bullying, Katherine found the 12-year-old girl. When Katherine told Ash, he uncovered the boys’ real names and figured out that they were high-schoolers in Abilene, Texas. Then he pieced together screenshots of their nasty tweets, along with their names and information about the schools they attended, and released it all in a public outing (called a “dox”). “I am sick of seeing people who think they can get away with breaking someone’s confidence and planting seeds of self-hate into someone’s head,” he wrote to them in the dox. “What gives you the fucking right to attack someone to such a breaking point? If you are vile enough to do so and stupid enough to do so on a public forum, such as a social website, then you should know this ... We will find you and we will highlight your despicable behaviour for all to see.”

“I informed them that the damage had been done and there was no going back,” he explained to me. “They understood this to be an act by Anonymous when they were then

messed in the hundreds.” At first the boys railed against Ash on Twitter, and one played down his involvement, denying that he had ever threatened to rape the girl. But after a while, two of the boys began sending remorseful messages. “For two solid days, every time we logged on, we had another apology from them,” Ash said. “You hear a lot of lies and fake apologies, and these guys seemed quite sincere.” Katherine thought the boys hadn’t understood what impact their tweets would have on the girl receiving them—they hadn’t thought of her as a real person. “They were actually shocked,” she said. “I’m sure they didn’t mean to actually rape a little girl. But she was *scared*. When they started to understand that, we started talking to them about anti-bullying initiatives they could bring to their schools.”

I tried contacting the four boys to ask what they made of their encounter with Anonymous, and I heard back from one of them. He said that at first, he thought the girl’s account was fake; then he assumed she wasn’t upset, because she didn’t block the messages he and the other boys were sending. Then Ash stepped in. “When i found out she was hurt by it i had felt horrible,” wrote to me in an e-mail. “I honestly don’t want to put anyone down. i just like to laugh and it was horrible to know just how hurt she was.” He also wrote, “It was shocking to see how big [Anonymous was] and what they do.”

Ash also e-mailed his catalog of the boys’ tweets to their principals and superintendents. I called the school officials and reached Joey Light, the superintendent for one of the districts in Abilene. He said that when Anonymous contacted him, “to be truthful, I didn’t know what it was. At first the whole thing seemed sketchy.” Along with the e-mails from Ash, Light got an anonymous phone call from a local number urging him to take action against the boys. Light turned over the materials Ash had gathered to the police officer stationed at the district’s high school, who established that one of the boys had been a student there.

The officer investigated, and determined that the boy hadn’t done anything to cause problems at school. That meant Light couldn’t punish him, he said. “I realize bullying takes a lot of forms, but our student couldn’t have harmed this girl physically in any way,” he continued. “If you can’t show a disruption at school, the courts tell us, that’s none of our business.” Still, Light told me he that he felt appreciative of Anonymous for intervening. “I don’t have the technical expertise or the time to keep track of every kid on Facebook or Twitter or whatever,” the superintendent said. “It was unusual, sure, but we would have never done anything if they hadn’t notified us.”

I talked with Ash and Katherine over Skype about a week after their Texas operation. I wanted to know how they’d conceived of the action they’d taken. Were they dispensing rough justice to one batch of heartless kids? Or were they trying to address cyberbullying more broadly, and if so, how?

Ash and Katherine said they'd seen lots of abuse of teenagers on social-networking sites, and most of the time, no adult seemed to know about it or intervene. They didn't blame the kids' parents for being clueless, but once they spotted danger, as they thought they had in this case, they couldn't bear to just stand by. "It sounds harsh to say we're teaching people a lesson, but they need to realize there are consequences for their actions," Ash said.

He and Katherine don't have professional experience working with teenagers, and I'm sure there are educators and parents who'd see them as suspect rather than helpful. But reading through the hate-filled tweets, I couldn't help thinking that justice Anonymous-style is better than no justice at all. In their own way, Ash and Katherine were stepping into the same breach that Henry Lieberman is trying to fill. And while sites like Facebook and Twitter are still working out ways to address harassment comprehensively, I find myself agreeing with Ash that "someone needs to teach these kids to be mindful, and anyone doing that is a good thing."

For Ash and Katherine, this has been the beginning of [#OpAntiBully](#), an operation that has a Twitter account providing resource lists and links to abuse-report forms. Depending on the case, Ash says, between 50 and 1,000 people—some of whom are part of Anonymous and some of whom are outside recruits—can come together to report an abusive user, or bombard him with angry tweets, or offer support to a target. "It's much more refined now," he told me over e-mail. "Certain people know the targets, and everyone contacts each other via DMs [direct messages]."

In a better online world, it wouldn't be up to Anonymous hackers to swoop in on behalf of vulnerable teenagers. But social networks still present tricky terrain for young people, with traps that other kids spring for them. My own view is that, as parents, we should demand more from these sites, by holding them accountable for enforcing their own rules. After all, collectively, we have consumer power here—along with our kids, we're the site's customers. And as Henry Lieberman's work at MIT demonstrates, it is feasible to take stronger action against cyberbullying. If Facebook and Twitter don't like his solution, surely they have the resources to come up with a few more of their own.

This article available online at:

<http://www.theatlantic.com/magazine/archive/2013/03/how-to-stop-bullies/309217/>

Copyright © 2015 by The Atlantic Monthly Group. All Rights Reserved.