THE CONVERSATION
Academic rigour, journalistic flair

Six things social media users and businesses can do to combat hate online

Published: March 7, 2022 10.05am GMT

Denitsa Dineva
Lecturer in Marketing and Strategy, Cardiff University

Lala Azizli/Unsplash

Online hostility has become a bigger problem over recent years, particularly with people spending more time on social media during the COVID-19 pandemic. A US survey found four in ten Americans have experienced harassment online – with three-quarters reporting that the most recent abuse happened on social media.

When online hostility happens on a continued basis it can be classified into a range of behaviours such as trolling, bullying and harassment.

More severe forms of online hostility can have real-world consequences for those affected, such as mental and emotional distress.

Debates about who should be responsible for the management of online hostility have been taking place over the last decade, but with little agreement. I would argue that three different sectors need to be involved: social media platforms, the companies that host business pages on social media, and users themselves.

The foundation of online hostility moderation lies with social media platforms. They must continuously update their processes and features to minimise the problem. We regularly hear that social media platforms are not doing enough to counter online hostility, and this may be true. In particular, I believe platforms could do more to educate companies and people about the available features designed to address hostility, and how to implement these appropriately.

What you can do

While social media platforms and businesses each play crucial roles in moderation, it's social media users who experience hostility first-hand, either as observers or victims.

There is no one-size-fits-all approach to responding to online hostility, but here are three courses of action you might consider.

1. Defend the victims

Providing support to the victims of hostility by challenging the aggressor and asking them to stop could be a viable option in less severe instances of online hostility. Recent research has shown that this can make the victim feel satisfied with the online brand community (for example, the Facebook fanpage) where the hostility occurred.

While this can be an effective way to combat hostility, and can make the victim feel supported, there's also a risk that it can escalate the situation, with the aggressor continuing to attack the victim, or attacking you. In this case, the two options below may be better.

Read more: Social media helps reveal people's racist views – so why don't tech firms do more to stop hate speech?

---

2. Hide, mute or block hostile content

Hiding, muting or blocking hostile content or users could be appropriate where users feel less comfortable to respond, but don't want to continue to be exposed to harmful content.

This isn't just for victims. We know harassment doesn't have to be experienced directly to be upsetting. This option puts the user in control of the situation and allows them to either temporarily or permanently block hostility (depending on whether it's a one-off or happening frequently).

3. Report hostile content

In instances of severe and repeated hostility, reporting content and users to companies or platforms is a suitable option. This requires the user to describe the incident and type of hostility that has occurred.

A woman looks at her smartphone, appears unhappy.

Experiences of online hostility can affect a person's mental wellbeing. Prostock-studio/Shutterstock

What businesses can do

Companies that manage social media pages can also block and report content and users, but they have other tools at their disposal, too.

For example, social media platforms enable companies to self-moderate their business pages by blocking offensive words from appearing. Businesses and brands that manage a Facebook page can choose up to 1,000 keywords to block in any language (these can include words, phrases and even emojis). If a user posts a comment containing one of the blocked words, their post will not be shown unless the page's administrator chooses to publish it.

While these tools may help to a degree, automated platform features alone are not enough. Technology is increasingly sophisticated, but it's difficult for machines to determine whether a particular comment or post is appropriate or not, regardless of the language used. Platforms also rely on human moderators, but these are a finite resource.

As part of my research into hostility moderation, I have looked at the different strategies which companies and brands are choosing to adopt. These include:

1. **Impartial or neutral strategies** mean the companies do not take a particular side during incidents, but provide further information on the topic at the root of the hostility.

2. **Cooperative moderation strategies** involve reinforcing positive comments and interactions by acknowledging those users who support others during incidents of hostility.

3. **Authoritative strategies** focus on moderating hostility by referring to the business page engagement rules and, in more extreme instances, by temporarily or permanently blocking users from posting comments.

My research has also found that an authoritative approach to moderation, in requesting users to interact in a more civil manner, generates the most positive attitudes towards the company, and a perception that it has a level of social responsibility.

Read more: What Facebook isn't telling us about its fight against online abuse

---

Ultimately, we all have a role to play to address hostility online. Social media platforms are not perfect, but they have made moderation tools widely available, and we should use them where it's warranted.